Fast Evaluation of Radial Basis Functions: Theory and Application

A thesis presented for the Degree of
Doctor of Philosophy in Mathematics
at the University of Canterbury by

J. B. Cherrie

University of Canterbury
Christchurch, New Zealand
May 30, 2000

Abstract

Radial Basis Functions (RBFs) have proven to be successful interpolants to scattered data. However, the perceived high computational costs for fitting and evaluating the RBFs associated with large data sets have hindered their application to many real world problems. This thesis is concerned with the "fast" evaluation of RBFs: the $\mathcal{O}(N^2)$ process of evaluation at all centres is reduced to $\mathcal{O}(N \log N)$ or even $\mathcal{O}(N)$. The required theory is developed for polyharmonic RBFs in 4-dimensions and for multiquadric RBFs in arbitrary dimensions. These methods are applied to fit surfaces to scattered data containing many tens of thousands of points.

Chapter 1 presents background material on the established theory of radial basis function interpolation and describes the how the Fast Multipole Method may used for the fast evaluation of radial basis functions.

Chapter 2 contains work published in the paper "Fast Evaluation of Radial Basis Functions: Methods for Four-Dimensional Polyharmonic Splines," co-authored by R. K. Beatson and D. L. Ragozin.

Chapter 3 contains work published in the paper "Fast Evaluation of Radial Basis Functions: Methods for Generalised Multiquadrics in $\mathbb{R}^n$," co-authored by R. K. Beatson and G. N. Newsam.

Chapter 4 contains work that has been done in conjunction with Applied Research Associates NZ Ltd.

Jon Cherrie, 29 May 2000.

Acknowledgements

First and foremost, I would like to thank my supervisor, Rick Beatson, for doing what supervisors do as well as those extra little things like lunches at Foo San.

My postgraduate studies could not have happened without the financial support of a University of Canterbury Doctoral Scholarship and the Shirtcliffe Fellowship, administered by the NZVCC. A big thanks to them.

I would like to thank Rick Fright and everyone else at Applied Research Associates NZ Ltd for inventing their very cool laser scanner and thus providing a number of research problems for applied mathematicians like me. I also wish to thank ARA for funding me through PGSF subcontract DRF601. Particular thanks go to Tim Mitchell and his implementation of the Domain Decomposition RBF fitter and to Tim Evans and Jonathan Carr for their isosurfacing programs *MCubes* and *FastSurf*.

Thanks to David Ragozin, my coauthor on the 4D polyharmonic paper [8], for all the assistance and ideas he provided.

The Department of Mathematics and Statistics has been very supportive. Special thanks to Ian Coope for his assistance on the optimization side of the Marching Triangles algorithm and for being the other United supporter in the department; Chris Hann, Andy McKenzie, Kahn Mason, Cameron Mouat, Paul Shorten, Matthew "Buddha" Wilkins and the other postgraduates in the department who let me bounce my ideas off them when they probably only had little idea, and even less interest, in what I was talking about and would really have much rather been doing something else.

The final form of this thesis was greatly enhanced by the critical comments of the examiners. I appreciate the time, effort and expertise they contributed.

Most especially I would like to thank Cressida Harding for all the support she has given me during both my undergraduate and postgraduate studies and for doing all of the cleaning, cooking and other domestic chores while I was writing this thesis. Cressida also acted as my writing adviser, correcting my grammar and generally improving the way this document reads. Thanks.

Contents

Chapter 1

Introduction

This introduction is a review of the theory of radial basis functions. Its purpose is to give statements, and often proofs, of the most fundamental known results. It also provides the necessary background for the original work of the later chapters.

A *radial basis function* (RBF) is a function of the form

$$s = p + \sum_{i=1}^{N} \lambda_i \phi(| \cdot - x_i|),$$

where p is a polynomial of low degree and the basic function ϕ is a real valued function on $[0, \infty)$ usually unbounded and of non-compact support (see, *e.g.*, Powell [54]). The points x_i are usually referred to as the centres of the RBF.

In a large scale comparison of methods for interpolating scattered data in two dimensions, Franke [27] identified radial basis functions as "often giving the most accurate results of all tested methods" and producing surfaces that are "usually pleasing and very smooth." Hardy [33] lists a large number of areas where RBFs have been employed, usually with better results than competing techniques. These applications include geodesy, geophysics, signal processing, and hydrology. RBFs have also been successfully employed for medical imaging [18] and morphing of surfaces in three dimensions [66]. However, despite their advantages perceived high computational costs for fitting and evaluating the RBFs associated with large data sets have hindered their application to many real world problems.

> The global interpolation methods with Duchon's "thin plate splines" and Hardy's multiquadrics are considered to be of high quality; however, their application is limited, due to computational difficulties, to ~150 data points. *Dyn, Levin & Rippa*, 1986 [24].

> Practical problems often arise with many more than 10,000 data sites; for example, in aeromagnetic survey work it is common to have 50,000 to 100,000 observations in a single data set. We believe that such problems will indefinitely remain beyond the scope of thin-plate splines. *Sibson & Stone*, 1991 [62].

> The most accurate results of registration of images with local distortions were obtained by using the surface spline mapping functions. As shown below, their direct use has extreme computing complexity and is not suitable for practical applications. *Flusser*, 1992 [26].

This thesis discusses how the computational difficulties of evaluating RBFs may be overcome by the use of "fast" evaluation techniques. It is shown that the $\mathcal{O}(N^2)$ process of evaluation at all centres is reduced to $\mathcal{O}(N \log N)$ or even $\mathcal{O}(N)$. Evaluation at all centres is an integral part an iterative fitting process and thus fast fitting requires fast evaluation.

Fast evaluation of RBFs is achieved via the Fast Multipole Method of Greengard & Rokhlin [29]. However, to implement this method for a given RBF a considerable amount of theoretical knowledge of the corresponding basic function must be known.

The remainder of this chapter backgrounds the interpolation problem and its solution via RBFs. It also describes the fast multipole method (FMM) and its application to the evaluation of RBFs.

1.1 Interpolation

The solution to a great many problems in a variety of areas depends on solving an interpolation problem.

Problem 1.1. *Given a set of distinct nodes $X = \{x_i\}_{i=1}^N \subset \mathbb{R}^d$ and a set of function values $\{f_i\}_{i=1}^N \subset \mathbb{R}$, find an* interpolant $s : \mathbb{R}^d \to \mathbb{R}$ *such that*

$$s(x_i) = f_i, \qquad i = 0, \ldots, N. \tag{1.1}$$

While this is not the most general form of the interpolation problem, it will suffice for needs of this dissertation. Throughout the rest of this thesis whenever a set of points in $\mathbb{R}^d$ is referred to as a set of nodes it is to be understood that these *nodes are distinct.* Note that d will used as the dimension of the Euclidean space that the nodes are sampled from.

Depending on the application, there may be further requirements placed on the interpolant. For example, it would be usual to require that the interpolant is at least continuous and often that the derivatives of the interpolant are also continuous up to some prescribed order.

One of the easiest ways to interpolate is to choose a set of basis functions $\{g_j\}_{j=1}^N$ and write the interpolant as a linear combination of them, *i.e.*,

$$s(x) = \sum_{j=0}^{N} \lambda_j g_j(x). \tag{1.2}$$

As most applications require a continuous interpolant, it will be assumed that the functions g_j are continuous. Assume that s is as given by Equation (1.2). Substitution of this form for s into Equation (1.1) leads to the linear system $\mathsf{A}\boldsymbol{\lambda} = \boldsymbol{f}$ to solve for the coefficients $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)^\mathsf{T}$, where

$$\mathsf{A} = \begin{bmatrix} g_1(x_1) & \cdots & g_N(x_1) \\ \vdots & \ddots & \vdots \\ g_1(x_N) & \cdots & g_N(x_N) \end{bmatrix} \tag{1.3}$$

and $\boldsymbol{f} = (f_1, \ldots, f_N)^\mathsf{T}$. Note that the matrix-vector product $\mathsf{A}\boldsymbol{\lambda}$ corresponds to evaluation of the spline s at all of the nodes. One consequence of this that will be of importance later is that this particular matrix-vector product may be calculated without forming or storing the matrix A, and possibly more quickly than by the naïve direct approach which requires $\mathcal{O}(N^2)$ flops.

The interpolant will be uniquely defined if $\det \mathsf{A} \neq 0$. This gives a test for the suitability of the basis functions given the node set X, but it does not give a means for choosing the functions. Although we might hope to choose a set of basis functions to give a unique interpolant for any node set, this is not possible when $d \geq 2$. As explained below, when $d \geq 2$, the functions g_j in Equation (1.2) must depend on the nodes otherwise the positions of those nodes will be restricted by the functions.

Let $G = \operatorname{span}\{g_j\}_{j=1}^N$ and let the functions g_j be defined on some prescribed domain $\Omega \subseteq \mathbb{R}^d$, *i.e.*, $G \subset C(\Omega)$. Properties such as existence and uniqueness of an interpolant of the form of that in Equation (1.2) are properties of the space G, not of the particular basis for G.

Definition 1.2. An N-dimensional vector space G of continuous functions on a domain Ω is a *Haar system* if zero is the only function in G with more than $N-1$ zeros in Ω.

This is equivalent to the condition $\det \mathsf{A} \neq 0$ for all choices of N distinct points $\{x_i\}_{i=1}^N \subset \Omega$, where the matrix A is defined by Equation (1.3) and the functions $\{g_j\}_{j=1}^N$ form a basis for G. Some authors use the term *Chebyshev system* instead of Haar System.

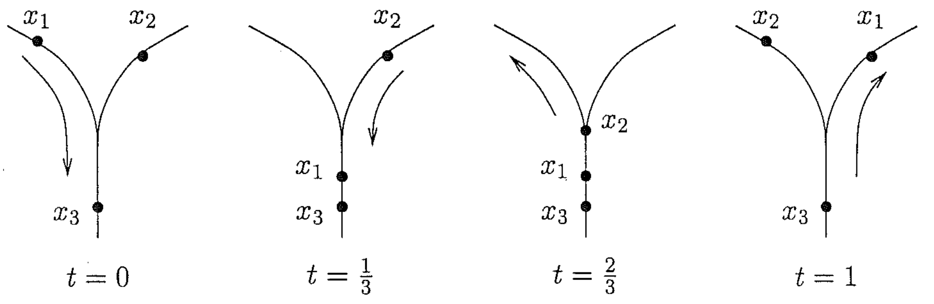


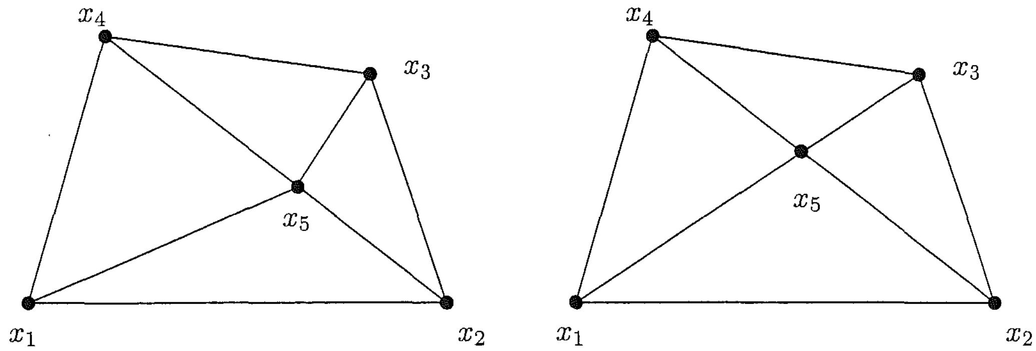
Figure 1.1: Standard example to show the ability to swap nodes in a "Y" shaped domain.

Assume that a set of basis functions, and thus a space G, have been chosen. If $G \subset C(\Omega)$ is to be a Haar system then there are very restricted choices for the domain Ω. Mairhuber [45] and McCullough & Wulbert [46], among others, showed that if $d \geq 2$ and any finite dimensional subspace of $C(\Omega)$ is to be a Haar system, then Ω must be homeomorphic to a subset of the unit circle. If Ω does not satisfy this condition then it is easily seen that G, as a subspace of $C(\Omega)$, cannot satisfy the Haar condition. For example, assume that Ω has a subset homeomorphic to a "Y" shape and consider the determinant of

$$\mathsf{A}(t) = \begin{bmatrix} g_1(x_1(t)) & g_2(x_1(t)) & g_3(x_1(t)) \\ g_1(x_2(t)) & g_2(x_2(t)) & g_3(x_2(t)) \\ g_1(x_3(t)) & g_2(x_3(t)) & g_3(x_3(t)) \end{bmatrix},$$

where the positions of the nodes x_1, x_2 and x_3 are continuous functions of some parameter t. Figure 1.1 shows that as t increases from 0 to 1 the nodes x_1 and x_2 may be moved in a continuous manner without ever being coincident (thus maintaining the distinctness condition) and thereby have their positions swapped. This process has the effect of swapping the first two rows of A and thus negating the determinant of A, *i.e.*, $\det \mathsf{A}(0) = -\det \mathsf{A}(1)$. As A, and thus $\det \mathsf{A}$, is a continuous function of t, the Intermediate Value Theorem shows that there is a $t^\star \in [0,1]$ such $\det \mathsf{A}(t^\star) = 0$. That is, for some configuration of x_1, x_2 and x_3, the corresponding interpolation system is non-invertible. This in turn implies that the restriction of G to this set of nodes has dimension less than 3. Hence $G \subset C(\Omega)$ cannot be a Haar system. A similar argument shows that Ω cannot contain a proper subset homeomorphic to the circle.

Another example of this dimension instability with respect to geometry is shown in Figure 1.2. Consider the dimension of the space $\mathcal{X}$ of bivariate continuously differentiable piecewise quadratics on a simple mesh defined by the triangulation of five knots. The knots are



(a) The fifth node x_5 lies on one of the diagonals but not the other.

(b) The fifth node x_5 is at the intersection of the two diagonals.

Figure 1.2: Sensitivity of dimension to node location.

arranged so that four of the knots are at the corners of a proper quadrilateral and the fifth is in the interior of quadrilateral. If the fifth knot lies on the intersection of the two diagonals of the quadrilateral then the dimension of $\mathcal{X}$ is 8. However, if the fifth knot lies anywhere else in the interior then the dimension of $\mathcal{X}$ is 7 [34, 60].

It follows from these arguments that it is not always possible to the solve the Interpolation Problem 1.1 for an arbitrary node set $\{x_i\}_{i=1}^N$ by considering linear combinations of fixed set of functions $\{g_i\}_{i=1}^N$. Thus to develop a method for interpolating on such node sets, the functions g_i must depend on the nodes. One of the simplest ways to achieve this dependence is to consider translates of a single function g, *i.e.*, the interpolant is chosen from span$\{g_i = g(\cdot - x_i)\}_{i=1}^N$. The following subsections present some possibilities for g such that the corresponding interpolation problem has a solution.

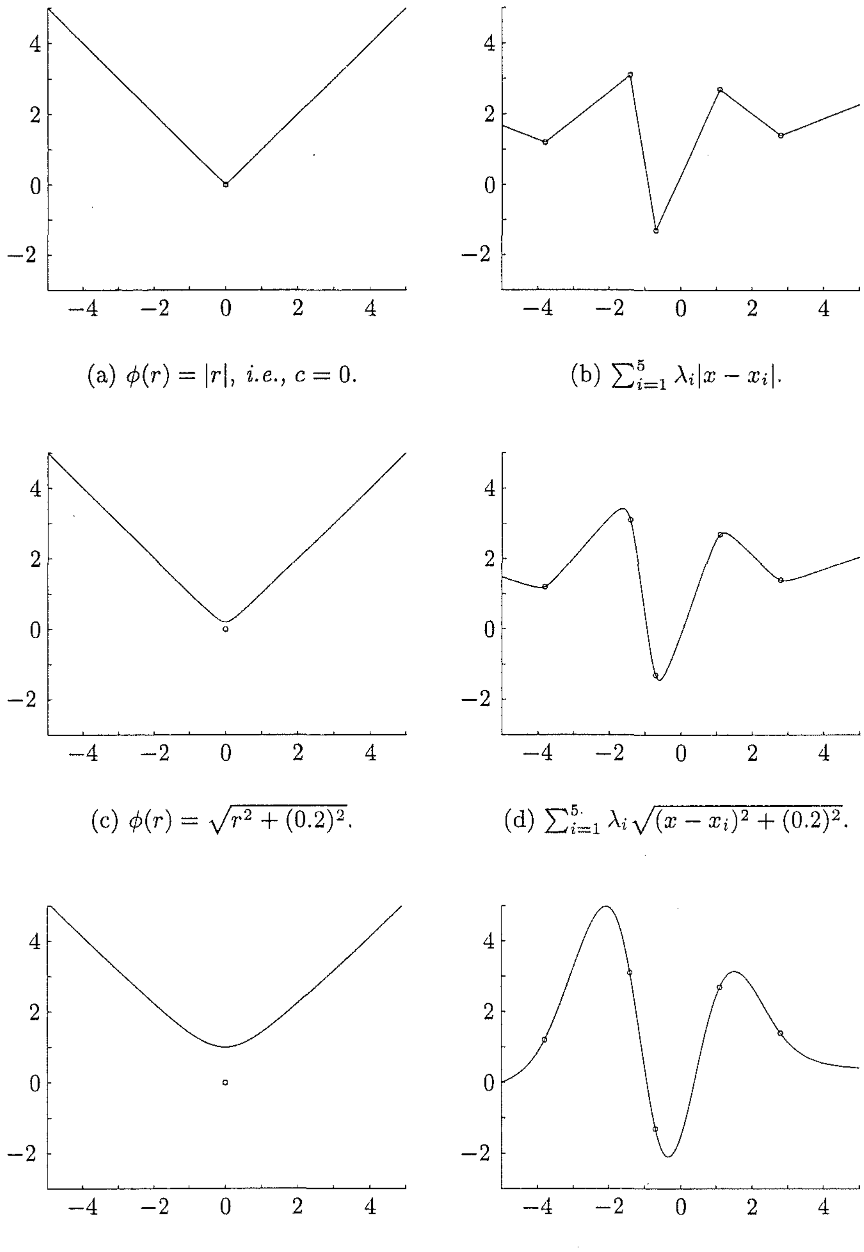
1.1.1 Multiquadrics

Multiquadrics were first introduced by Hardy [32, 33] as the result of a trial and error process which attempted to deal with "*the frustrations of trying to use various harmonic and polynomial series to represent topography from relatively few data points.*" His solution was to use the basis functions

$$g_j(x) = \sqrt{|x - x_j|^2 + c^2},$$

where c is some real constant. These were later generalized to

$$g_j(x) = \left(|x - x_j|^2 + c^2\right)^{m/2},$$



(a) $\phi(r) = |r|$, *i.e.*, $c = 0$.

(b) $\sum_{i=1}^{5} \lambda_i |x - x_i|$.

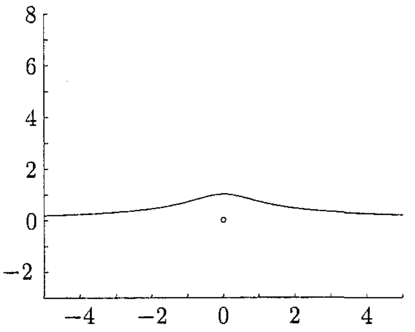
(c) $\phi(r) = \sqrt{r^2 + (0.2)^2}$.

(d) $\sum_{i=1}^{5} \lambda_i \sqrt{(x - x_i)^2 + (0.2)^2}$.

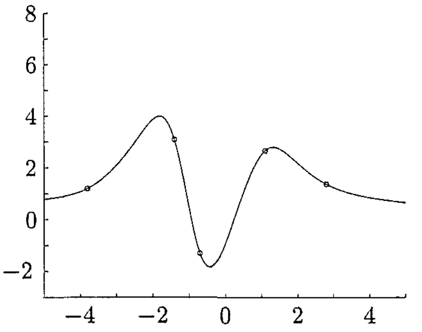
(e) $\phi(r) = \sqrt{r^2 + (1.0)^2}$.

(f) $\sum_{i=1}^{5} \lambda_i \sqrt{(x - x_i)^2 + (1.0)^2}$.

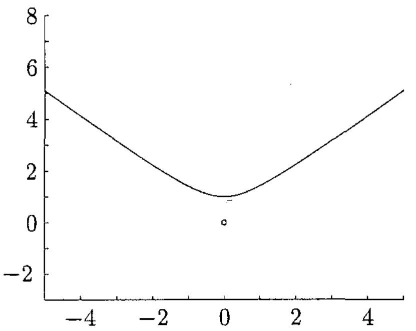
Figure 1.3: Multiquadric basic functions (left) and splines (right) that interpolate the data in Table 1.1.



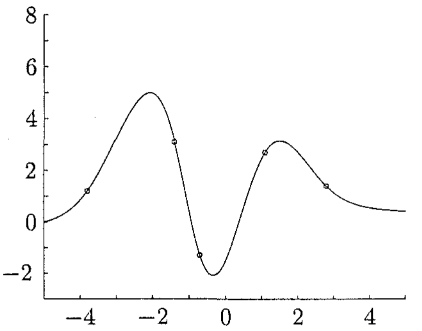
(a) Inverse multiquadric, $\phi(r) = (r^2 + 1.0)^{-1/2}$.



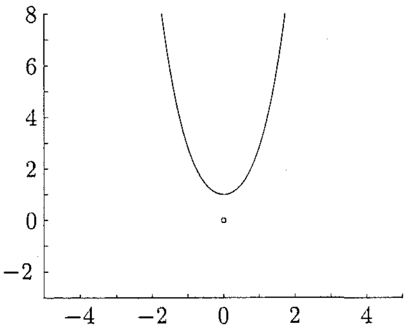
(b) $\sum_{i=1}^{5} \lambda_i \left((x - x_i)^2 + 1.0\right)^{-1/2}$.



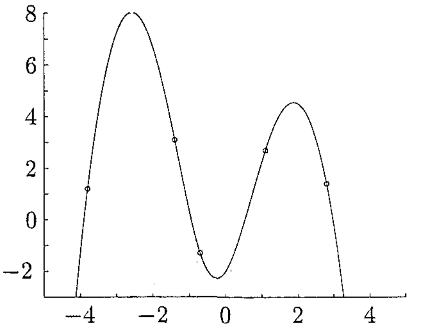
(c) The ordinary multiquadric, $\phi(r) = (r^2 + 1.0)^{1/2}$.



(d) $\sum_{i=1}^{5} \lambda_i \left((x - x_i)^2 + 1.0\right)^{1/2}$.



(e) A generalized multiquadric, $\phi(r) = (r^2 + 1.0)^{3/2}$.



(f) $\sum_{i=1}^{5} \lambda_i \left((x - x_i)^2 + 1.0\right)^{3/2}$.

Figure 1.4: Generalised multiquadric basic functions (left) and splines (right) that interpolate the data in Table 1.1. The constant in each example is $c = 1.0$.

where m is an odd integer. If m is even and positive then this reduces the interpolant to a polynomial. The example of Figure 1.1 does not apply to the space spanned by these functions since the basis functions change with the movement of the nodes.

Notice that a particular basis function g_j is rotation invariant for rotations about x_j, *i.e.*, g_j is of the form

$$g_j = \phi(|\cdot - x_j|),$$

where ϕ is a function defined on $[0, \infty)$. This will prove to be very important in ensuring that the corresponding interpolation system is invertible for a large class of node sets.

Figures 1.3 and 1.4 show some examples of the multiquadric function used to interpolate a simple 1D data set [19, p. 1]. The graphs on the left show a single basis function, while those on the right show a linear combination of five basis functions used to interpolate some data. The data used for each interpolation is given in the following table.

x	-3.8	-1.4	-0.7	1.1	2.8
y	1.2	3.1	-1.3	2.7	1.4

Table 1.1: Data for example interpolations.

1.1.2 Quadratic penalties and thin-plate splines

This subsection concerns interpolants that minimize some suitable energy semi-norm. This problem was first solved by Duchon [22] and Meinguet [47, 48]. However, the presentation here is based on the approach of Golomb & Weinberger [28] and Light [43].

Firstly, a few definitions will be required. For the multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$, the αth derivative is defined as

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial\xi_1^{\alpha_1} \cdots \partial\xi_d^{\alpha_d}} = \left(\frac{\partial}{\partial\xi_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial\xi_d}\right)^{\alpha_d}$$

where $|\alpha| = |\alpha_1| + \cdots + |\alpha_d|$ and where $\xi_1, \ldots, \xi_d$ are the components of vectors in $\mathbb{R}^d$. The binomial symbol is generalized to give the multinomial symbol

$$\binom{m}{\alpha} = \frac{m!}{\alpha_1! \cdots \alpha_d!}.$$

Let $\mathcal{D}(\mathbb{R}^d)$ be the space of test functions, *i.e.*, the space of infinitely differentiable, compactly supported functions defined on $\mathbb{R}^d$. The interpolant will be chosen from a subspace of

$\mathcal{D}'(\mathbb{R}^d)$, the space of *distributions* or functionals on $\mathcal{D}(\mathbb{R}^d)$. Specifically, the interpolant will be chosen from the Beppo-Levi type space,

$$\mathcal{X} = \mathrm{BL}^{(m)}(\mathbb{R}^d) = \left\{ s \in \mathcal{D}'(\mathbb{R}^d) : D^\alpha s \in L_2(\mathbb{R}^d), \text{ for all } |\alpha| = m \right\}. \tag{1.4}$$

As will be shown below, to ensure that $\mathcal{X}$ consists of functions, it will be necessary to impose the restriction $m > d/2$. On this new space, define the semi-inner product

$$(s, t)_m = \int_{\mathbb{R}^d} \sum_{|\alpha|=m} \binom{m}{\alpha} D^\alpha s(x) D^\alpha t(x) dx, \tag{1.5}$$

and denote the semi-norm it induces on $\mathcal{X}$ by $|\cdot|_m$. Note that this semi-norm is rotation invariant in the sense that it is invariant under rotation of the underlying coordinate system. The kernel or null space of this semi-norm is $\pi_{m-1} = \pi_{m-1}(\mathbb{R}^d)$, which is the space of polynomials on $\mathbb{R}^d$ of degree at most $m-1$.

There is now enough machinery to state the problem that is the subject of this subsection.

Problem 1.3. *Given a set of nodes $X = \{x_i\}_{i=1}^N \subset \mathbb{R}^d$ and a set of function values $\{f_i\}_{i=1}^N \subset \mathbb{R}$, find $s^\star \in S$ such that*

$$|s^\star|_m = \inf_{s \in S} |s|_m, \tag{1.6a}$$

where S is the affine space of interpolants to data $\{f_i\}$ on the node set X, i.e.,

$$S = \{ s \in \mathcal{X} : s(x_i) = f_i, \quad i = 1, \ldots, N \}. \tag{1.6b}$$

The solution, $s^\star$, to this problem is referred to as the minimal norm interpolant to the data $\{(x_i, f_i)\}_{i=1}^N$. This problem is easily adapted to interpolation on spheres. See, for example, Levesley, *et al.* [42] or Wahba [68].

A useful way to think of the semi-norm $|\cdot|_m$ is as a measure of energy. For example, the classic thin-plate spline interpolant on $\mathbb{R}^2$ is the interpolating surface that minimizes the bending energy of a thin metal plate of infinite extent (see, Harder & Desmarais [31]).

The following theorem is very important in allowing a number of results to be brought to bear on Problem 1.3.

Theorem 1.4. (Meinguet [47, Theorem 1]) *Suppose $m > d/2$. Then the semi-normed space $\mathcal{X}$, defined by Equations* (1.4) *and* (1.5)*, is a semi-Hilbert function space of continuous functions on $\mathbb{R}^d$. Furthermore, all of the evaluation linear functionals on $\mathcal{X}$ with finite support in $\mathbb{R}^d$ that annihilate π_{m-1} are bounded.*

For there to be a unique solution to Problem 1.3 there will need to be a mild condition on the location of the interpolation nodes.

Definition 1.5. A finite set $X \subset \mathbb{R}^d$ is a π_{m-1}-*unisolvent* set if the only polynomial $p \in \pi_{m-1}$ that satisfies

$$p(x_i) = 0 \qquad \text{for all } x_i \in X$$

is the zero polynomial.

Example 1.6. If X is non-empty, then X is π_0-unisolvent. If all the points in X do not lie on a single hyperplane in $\mathbb{R}^d$ then X is π_1-unisolvent.

Let $\ell_m = \dim(\pi_{m-1})$ and order the nodes so that $\{x_i\}_{i=1}^{\ell_m}$ is a π_{m-1}-unisolvent set. Define an inner product on $\mathcal{X}$,

$$\langle s, t\rangle_m = (s, t)_m + \sum_{i=1}^{\ell_m} s(x_i)t(x_i). \tag{1.7}$$

and denote the norm associated with this inner product by $\|\cdot\|_m$. Using the fact that $\mathcal{X}$ with $(\cdot,\cdot)$ is a semi-Hilbert space, it is easy to show that $\mathcal{X}$ with $\langle\cdot,\cdot\rangle$ is a Hilbert space. Note that $s^\star \in S$ minimizes $|\cdot|_m$ over S if and only if it minimizes $\|\cdot\|_m$ over S.

Theorem 1.8, which in some sense solves Problem 1.3, is a direct consequence of the following classical result and the fact that $\mathcal{X}$ is a Hilbert space.

Lemma 1.7. *Let H be an inner product space and G be a subspace of H. If $f \in H$ then $g^\star$ is a best approximation to f from G if and only if*

$$\langle f - g^\star, g\rangle = 0, \qquad \textit{for all } g \in G.$$

Let G be an affine space in H. Now take an arbitrary but fixed element $\widetilde{g} \in G$ and shift G by $-\widetilde{g}$ so that $\widetilde{G} = G - \widetilde{g}$ is a subspace of H (see Figure 1.5). Consider the best approximation $\widetilde{g}^\star$ to $\widetilde{f} = f - \widetilde{g}$ from $\widetilde{G}$. Lemma 1.7 tells us that

$$\langle \widetilde{f} - \widetilde{g}^\star, h\rangle = 0, \qquad \text{for all } h \in \widetilde{G}.$$

For any $g \in G$ there is an $h \in \widetilde{G}$ such that $g = h + \widetilde{g}$. Furthermore, since $\widetilde{g}^\star \in \widetilde{G}$, there is a unique $g^\star \in G$ such that $\widetilde{g}^\star = g^\star + \widetilde{g}$. It now follows that $g^\star$ is the best approximation to f from G if and only if

$$\langle f - g^\star, g - \widetilde{g}\rangle = 0, \qquad \text{for all } g, \widetilde{g} \in G.$$

An interpretation of this is that the 'error' $f - g^\star$ is orthogonal to any feasible 'correction' that might be made to $g^\star$.

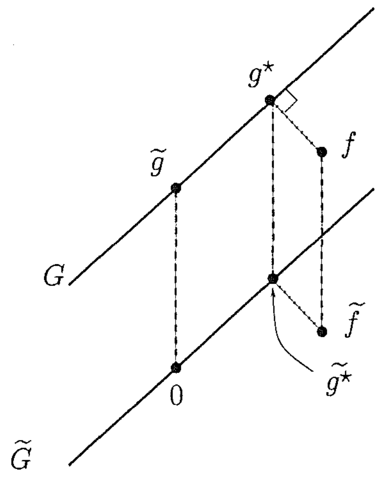


Figure 1.5: Best approximation from an affine space.

Since $\mathcal{X}$ is a Hilbert space of functions on $\mathbb{R}^d$, the Riesz representation theorem implies that for each $x \in \mathbb{R}^d$ there is a *representer* $x^* \in \mathcal{X}$ such that

$$s(x) = \langle x^*, s \rangle, \qquad \text{for all } s \in \mathcal{X}.$$

Theorem 1.8. (Light [43, Theorem 1.1]) *Let X be a π_{m-1}-unisolvent set in $\mathbb{R}^d$. Let $s^\star$ be the element of minimal norm in the set of interpolants*

$$S = \{v \in \mathcal{X} : \langle x_i^*, v\rangle_m = f_i, \ i = 1, \dots, N\}$$

to the data $\{(x_i, f_i)\}$. Then

$$s^\star = \sum_{j=1}^{N} \nu_j x_j^*$$

where the coefficients ν_i are determined by the equations

$$s^\star(x_i) = \langle x_i^*, s^\star\rangle_m = \sum_{j=1}^{N} \nu_j \langle x_i^*, x_j^*\rangle_m = f_i, \qquad i = 1, \dots, N. \tag{1.8}$$

Proof. First the existence of such an $s^\star$ will be proved. Let $\boldsymbol{\nu} = (\nu_1, \dots, \nu_N)^\mathsf{T}$ be an arbitrary vector of coefficients. Then

$$0 \le \left\| \sum_{j=1}^{N} \nu_j x_j^* \right\|_m^2 = \left\langle \sum_{i=1}^{N} \nu_i x_i^*, \sum_{j=1}^{N} \nu_j x_j^* \right\rangle_m = \sum_{i,j=1}^{N} \nu_i \nu_j \left\langle x_i^*, x_j^* \right\rangle_m,$$

and, because of the linear independence of the evaluation functionals, there is equality if and only if $\boldsymbol{\nu} = 0$. Hence the matrix $\mathsf{A}_{i,j} = \langle x_i^*, x_j^*\rangle_m$ is positive definite and existence of $s^\star$ is proved.

The problem of minimizing the norm over S can be thought of as finding the best approximation to 0 from S. Thus the 'error' $s^\star - 0$ is orthogonal to $s^\star - s$ for all $s \in S$. That is, $s^\star$ is orthogonal to

$$\{s \in \mathcal{X} : \langle x_i^*, s\rangle_m = 0,\ i = 1, \dots, N\} = \bigcap_{i=1}^{N} x_i^{*\perp},$$

where $\cdot^\perp$ denotes the orthogonal complement in $\mathcal{X}$ with respect to the inner product (1.7). Since

$$\left(\bigcap_{i=1}^{N} x_i^{*\perp}\right)^\perp = \operatorname{span}\{x_i^*,\ i = 1, \dots, N\},$$

it now follows that $s^\star$ is of the form $s^\star = \sum_{j=1}^{N} \nu_i x_j^*$ and the requirement that $s^\star \in S$ leads to equations (1.8). ∎

The theorem is only of practical importance if we can find an explicit form for the representers $x^* \in \mathcal{X}$ for points $x \in \mathbb{R}^d$. Because of the form of the inner product (1.7), the points $\{x_i\}_{i=1}^{\ell_m}$ have particularly simple and nice representers in $\mathcal{X}$.

Lemma 1.9. (Light [43, Theorem 2.1]) *Let $\{x_i\}_{i=1}^{\ell_m} \subset \mathbb{R}^d$ be a π_{m-1}-unisolvent set and let $\{p_i\}_{i=1}^{\ell_m}$ be the Lagrange basis for π_{m-1} with respect to $\{x_i\}_{i=1}^{\ell_m}$,* i.e., *$p_j(x_i) = \delta_{i,j}$. Then p_i is the representer in $\mathcal{X}$ for x_i.*

Proof. Since π_{m-1} is the kernel for $(\cdot,\cdot)_m$,

$$(s,p)_m = 0, \qquad \text{for all } s \in \mathcal{X} \text{ and all } p \in \pi_{m-1}.$$

It now follows that

$$\langle s, p_i\rangle_m = (s, p_i)_m + \sum_{j=1}^{\ell_m} s(x_j)p_i(x_j) = s(x_i),$$

for $i = 1, \dots, N$ and for any $s \in \mathcal{X}$. ∎

Again, the special form of the inner product (1.7) can be used to find the representers in $\mathcal{X}$ for $x \in \mathbb{R}^d \backslash \{x_i\}_{i=1}^{\ell_m}$. Let Γ_0 be the orthogonal complement of π_{m-1} in $\mathcal{X}$ with respect to the inner product (1.7). Then,

$$\Gamma_0 = \{s \in \mathcal{X} : s(x_i) = 0,\ i = 1, \dots, \ell_m\}.$$

Since Γ_0 is a subspace of H it is a Hilbert space in its own right and thus for each $x \in \mathbb{R}^d$ there is a representer $g_x \in \Gamma_0$. Assume for the moment the representer $g_x \in \Gamma_0$ for $x \in \mathbb{R}^d \backslash \{x_i\}_{i=1}^{\ell_m}$ is known, *i.e.*, $g_x \in \Gamma_0$ is known such that

$$\langle \gamma, g_x\rangle_m = \gamma(x) \qquad \text{for all } \gamma \in \Gamma_0.$$

This representer and those of Lemma 1.9 can be used to calculate x^* for $x \in \mathbb{R}^d \backslash \{x_i\}_{i=1}^{\ell_m}$. Let $\mathcal{P} : \mathcal{X} \to \pi_{m-1}$ be the orthogonal projection

$$\mathcal{P}s = \sum_{i=1}^{\ell_m} s(x_i)p_i, \tag{1.9}$$

where $\{p_i\}_{i=1}^{\ell_m}$ is the Lagrange basis for π_{m-1} with respect to $\{x_i\}_{i=1}^{\ell_m}$. Note that $\big(\mathcal{P}f\big)(x_i) = f(x_i)$ and so $\mathcal{P}f$ is actually the interpolating polynomial on $\{x_i\}_{i=1}^{\ell_m}$. The significance of this fact is that $(I - \mathcal{P})s \in \Gamma_0$ for all $s \in \mathcal{X}$. Thus, for any $s \in \mathcal{X}$,

$$\langle s, g_x \rangle_m = \big\langle (I - \mathcal{P})s, g_x \big\rangle_m + \langle \mathcal{P}s, g_x \rangle_m = \big((I - \mathcal{P})s\big)(x) = s(x) - \sum_{i=1}^{\ell_m} s(x_i)p_i(x). \tag{1.10}$$

By Lemma 1.9 $s(x_i) = \langle s, p_i \rangle_m$, so that rearranging Equation (1.10) gives

$$s(x) = \langle s, g_x \rangle_m + \sum_{i=1}^{\ell_m} \langle s, p_i \rangle_m p_i(x) = \left\langle s, g_x + \sum_{i=1}^{\ell_m} p_i(x)p_i \right\rangle_m,$$

and it follows that

$$x^* = g_x + \sum_{i=1}^{\ell_m} p_i(x)p_i. \tag{1.11}$$

Therefore, to use Theorem 1.8 to solve Problem 1.3, all that is required is to find the representer g_x.

For any $\gamma \in \Gamma_0$ and any $s \in \mathcal{X}$,

$$\langle \gamma, s \rangle_m = (\gamma, s)_m + \sum_{i=1}^{\ell_m} \gamma(x_i)s(x_i) = (\gamma, s)_m,$$

since $\gamma(x_i) = 0$ for $i = 1, \ldots, \ell_m$. Using $[f, \phi]$ to denote the action of a distribution $f \in \mathcal{D}'$ on a test function $\phi \in \mathcal{D}$ allows us to use the distributional definition of the derivative. Thus, for any $\gamma \in \Gamma_0$,

$$\begin{aligned}
[\delta(\cdot - x), \gamma] = \gamma(x) = \langle \gamma, g_x \rangle_m &= (\gamma, g_x)_m \\
&= \int_{\mathbb{R}^d} \sum_{|\alpha|=m} \binom{m}{\alpha} D^\alpha g_x(y) D^\alpha \gamma(y) dy \\
&= \sum_{|\alpha|=m} \binom{m}{\alpha} [D^\alpha g_x, D^\alpha \gamma] \\
&= \left[(-1)^m \sum_{|\alpha|=m} \binom{m}{\alpha} D^{2\alpha} g_x, \gamma \right] \\
&= [(-1)^m \Delta^m g_x, \gamma] .
\end{aligned} \tag{1.12}$$

Now let $s \in \mathcal{X}$ and let $\mathcal{P}$ be the orthogonal projection (1.9). Because of the interpolatory nature of $\mathcal{P}$, $(I-\mathcal{P})s \in \Gamma_0$ and replacing γ in Equation 1.12 by $(I-\mathcal{P})s$ gives

$$\left[\delta(\cdot - x) - \sum_{i=1}^{\ell_m} p_i(x)\delta(\cdot - x_i), s\right] = [(-1)^m \Delta^m g_x, s].$$

for all $s \in \mathcal{X}$. This implies that g_x satisfies the distributional equation

$$(-1)^m \Delta^m g_x = \delta(\cdot - x) - \sum_{i=1}^{\ell_m} p_i(x)\delta(\cdot - x_i). \tag{1.13}$$

Let $E \in \mathcal{D}'$ be the fundamental solution of the iterated Laplacian, *i.e.*,

$$\Delta^m E = \sum_{|\alpha|=m} \binom{m}{\alpha} D^{2\alpha} E = \delta,$$

where δ is the Dirac measure. Then [61, p. 288],

$$E(x) = E_m^{(d)}(x) = \begin{cases} c_{m,d}|x|^{2m-d} \ln |x|, & \text{if } 2m \geq d \text{ and } d \text{ is even,} \\ c_{m,d}|x|^{2m-d}, & \text{otherwise,} \end{cases} \tag{1.14}$$

where

$$c_{m,d} = \begin{cases} \dfrac{(-1)^{d/2+1}}{2^{2m-1}\pi^{d/2}(m-1)!(m-d/2)!}, & \text{if } 2m \geq d \text{ and } d \text{ is even,} \\ \dfrac{(-1)^m \Gamma(d/2-m)}{2^{2m}\pi^{d/2}(m-1)!}, & \text{otherwise.} \end{cases}$$

Hence it follows that a solution, g_x, to Equation (1.13) is

$$\widetilde{g_x} = (-1)^m \left(E(\cdot - x) - \sum_{i=1}^{\ell_m} p_i(x) E(\cdot - x_i) \right).$$

However, $\widetilde{g_x} \notin \Gamma_0$. Since $\Delta^m p = 0$ for any $p \in \pi_{m-1}$, $\widetilde{g_x} + p$ is another solution to Equation (1.13). Thus $g_x = \widetilde{g_x} - \mathcal{P}\widetilde{g_x}$ is a solution to Equation (1.13) and, because of the interpolatory properties of $\mathcal{P}$, is a member of Γ_0. That is,

$$\begin{aligned} g_x &= (-1)^m (I-\mathcal{P}) \left(E(\cdot - x) - \sum_{i=1}^{\ell_m} p_i(x) E(\cdot - x_i) \right) \\ &= (-1)^m \left(E(\cdot - x) - \sum_{i=1}^{\ell_m} p_i(x) E(\cdot - x_i) - \sum_{i=1}^{\ell_m} E(x_i - x) p_i + \sum_{i,j=1}^{\ell_m} E(x_j - x_i) p_i(x) p_j \right) \end{aligned} \tag{1.15}$$

is the representer in Γ_0 for $x \in \mathbb{R}^d$. In an attempt to make the symmetry in g_x as clear as possible, the symbols E and g will be overloaded so that $E(x,y) = E(x-y)$ and $g(x,y) = g_x(y)$. Thus

$$g(x,y) = (-1)^m \Bigg(E(y-x) - \sum_{i=1}^{\ell_m} E(y-x_i)p_i(x) \\ - \sum_{i=1}^{\ell_m} E(x_i - x)p_i(y) + \sum_{i,j=1}^{\ell_m} E(x_j - x_i)p_i(x)p_j(y) \Bigg).$$

Now that the representers in $\mathcal{X}$ are known for all $x \in \mathbb{R}^d$ Theorem 1.8 may be used to solve Problem 1.3. Special note should be taken of the simple form of $s^\star$ in Theorem 1.10; it contains no function more complicated than the logarithm and each term is radially symmetric about the corresponding node x_i.

Theorem 1.10. *If $m > d/2$ and X is a π_{m-1}-unisolvent set, then the solution to Problem 1.3 has the form*

$$s^\star = p + \sum_{i=0}^{N} \lambda_i E_m^{(d)}(\cdot - x_i) \tag{1.16}$$

where p is a polynomial of degree $m-1$ or less and the coefficients $\{\lambda_i\}_{i=1}^N$ satisfy the orthogonality conditions

$$\sum_{i=1}^{N} \lambda_i q(x_i) = 0 \qquad \text{for all } q \in \pi_{m-1}. \tag{1.17}$$

Proof. By Theorem 1.8, the solution to Problem 1.3 is given by

$$s^\star = \sum_{j=1}^{N} \nu_j x_j^* = \sum_{j=1}^{\ell_m} \nu_j p_j + \sum_{j=\ell_m+1}^{N} \nu_j g_{x_j},$$

where the coefficients $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_N)^\top$ are given by Equation (1.8). As can be seen from Equation (1.15), the representer g_x is composed of a sum of shifts of E plus a polynomial of degree not exceeding $m-1$. Hence it follows that $s^\star$ may be similarly decomposed, *i.e.*,

$$s^\star = p + \sum_{i=0}^{N} \lambda_i E_m^{(d)}(\cdot - x_i),$$

where p is a polynomial of degree $m-1$ or less. However, there are some conditions on the coefficients λ_i. From Equation (1.15), the non-polynomial part of $s^\star$ is given by

$$\sum_{i=0}^{N} \lambda_i E(\cdot - x_i) = (-1)^m \sum_{j=\ell_m+1}^{N} \nu_j \left(E(\cdot - x_j) - \sum_{i=1}^{\ell_m} p_i(x_j) E(\cdot - x_i) \right) \\ = (-1)^{m+1} \sum_{i=1}^{\ell_m} \left(\sum_{j=\ell_m+1}^{N} \nu_j p_i(x_j) \right) E(\cdot - x_i) + (-1)^m \sum_{i=\ell_m+1}^{N} \nu_i E(\cdot - x_i).$$

Hence

$$\lambda_i = \begin{cases} (-1)^{m+1} \displaystyle\sum_{j=\ell_m+1}^{N} \nu_j p_i(x_j), & 1 \le i \le \ell_m, \\ (-1)^m \nu_i, & \ell_m + 1 \le i \le N, \end{cases}$$

or, if

$$\mathsf{Q} = (-1)^m \begin{bmatrix} -p_1(x_{\ell_m+1}) & -p_1(x_{\ell_m+2}) & \cdots & -p_1(x_N) \\ -p_2(x_{\ell_m+1}) & -p_2(x_{\ell_m+2}) & \cdots & -p_2(x_N) \\ \vdots & & \ddots & \\ -p_{\ell_m}(x_{\ell_m+1}) & -p_{\ell_m}(x_{\ell_m+2}) & \cdots & -p_{\ell_m}(x_N) \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

then $\boldsymbol{\lambda} = \mathsf{Q}\boldsymbol{\nu}$ gives the coefficients $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)^{\mathsf{T}}$ in terms of the coefficients $\boldsymbol{\nu} = (\nu_{\ell_m+1}, \ldots, \nu_N)^{\mathsf{T}}$. In other words, $\boldsymbol{\lambda}$ must be in the column space of Q. This condition is equivalent to the condition that $\boldsymbol{\lambda}$ be in the null space of P^{T} where P is any $N \times \ell_m$ (real valued) matrix whose rows span the null space of Q. Since $\{p_j\}_{j=1}^{\ell_m}$ is a Lagrange basis for $\pi_{m-1}(\mathbb{R}^d)$ with respect to the points $\{x_i\}_{i=1}^{\ell_m}$, one such matrix is given by $\mathsf{P}_{i,j} = p_j(x_i)$. The condition $\mathsf{P}^{\mathsf{T}}\boldsymbol{\lambda} = 0$ is equivalent to Equation (1.17). ∎

It should be noted that a general function of the form (1.16), *i.e.*, one without the conditions (1.17) imposed on the coefficients, is not a member of $\mathcal{X}$ as the mth derivative is not guaranteed to be integrable.

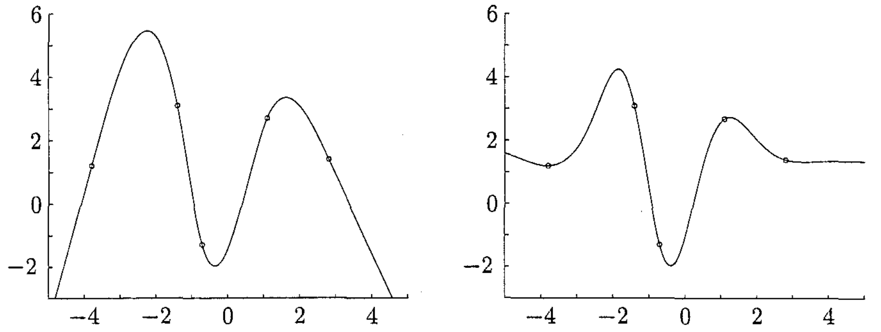
Example 1.11. The semi-norm

$$|s|_2 = \left(\int_{-\infty}^{\infty} \left(s''(y)\right)^2 dy \right)^{1/2}.$$

corresponds to the natural cubic spline of one variable ($m = 2$, $d = 1$). The minimum norm interpolant is given by

$$s^\star = p + \sum_{i=1}^{N} \lambda_i |\cdot - x_i|^3,$$

where p is a linear polynomial. Consider the natural cubic spline $s^\star$ and Gaussian spline, s_G, interpolants to the data in Table 1.1. The natural spline interpolant to this data has



(a) Natural cubic spline. (b) Gaussian spline.

Figure 1.6: Natural spline versus Gaussian spline.

a semi-norm of $|s^\star|_2 = 192.757$, while the Gaussian spline interpolant has a semi-norm of $|s_G|_2 = 269.565$. These two interpolants are shown in Figure 1.6.

Example 1.12. The classic thin-plate spline is the case when $d = 2$ and $m = 2$. Let $x = (\xi, \eta) \in \mathbb{R}^2$. Then the semi-norm is given by

$$|s|_2 = \left(\int_{\mathbb{R}^2} \left(\frac{\partial^2}{\partial \xi^2} s(x) \right)^2 + 2 \left(\frac{\partial^2}{\partial \xi \partial \eta} s(x) \right)^2 + \left(\frac{\partial^2}{\partial \eta^2} s(x) \right)^2 dx \right)^{1/2}$$

and the corresponding solution to Problem 1.3 is

$$s^\star(x) = p(x) + \sum_{i=0}^{n} \lambda_i |x - x_i|^2 \log |x - x_i|,$$

where p is a linear polynomial. This thin plate spline has proven a very useful general purpose interpolator for two dimensional data.

We conclude this subsection with a simple calculation of the semi-norm $|s^\star|_m$. Writing the semi-inner product in distributional form and using the distributional definition of the derivative,

$$\begin{aligned} |s^\star|_m^2 = \sum_{|\alpha|=m} \binom{m}{\alpha} [D^\alpha s^\star, D^\alpha s^\star] &= (-1)^m \left[\sum_{|\alpha|=m} \binom{m}{\alpha} D^{2\alpha} s^\star, s^\star \right] \\ &= (-1)^m \sum_{i=1}^{N} \lambda_i \left[\delta(\cdot - x_i), s^\star \right] \\ &= (-1)^m \sum_{i=1}^{N} \lambda_i s^\star(x_i) \\ &= (-1)^m \sum_{i,j=1}^{N} \lambda_i \lambda_j E(x_j - x_i) + \sum_{i=1}^{N} \lambda_i p(x_i) \end{aligned}$$

$$= (-1)^m \sum_{i,j=1}^{N} \lambda_i \lambda_j E(x_j - x_i),$$

since $\sum_{i=1}^{N} \lambda_i p(x_i) = 0$ for all polynomials p of degree $m-1$ or less.

1.1.3 Radial basis functions and conditional positive definiteness

Both the thin-plate splines of Duchon and the multiquadrics of Hardy are examples of radial basis functions.

Definition 1.13. A *radial basis function* (RBF) is a function of the form

$$s = p + \sum_{i=1}^{N} \lambda_i \phi(|\cdot - x_i|), \tag{1.18}$$

where p is a polynomial of low degree and ϕ is a real valued function on $[0, \infty)$.

This subsection concerns the simple conditions that ensure that the existence and uniqueness of an RBF interpolant. These conditions will be obtained using the theory of conditionally positive definite functions.

See Table 1.2 for a list of common choices for the basic function ϕ. Of the basic functions in this table, this thesis is principally concerned with the polyharmonic (including the linear, cubic, and thin-plate spline) and multiquadric basic functions.

The definition of an RBF may be generalized by saying that a radial basis function has the form

$$s = p + \sum_{i=1}^{N} \lambda_i \Phi(\cdot - x_i),$$

where Φ is any even function mapping $\mathbb{R}^d$ to $\mathbb{R}$. Schaback [58, 59] gives related results for this more general setting. However, the definition as given will be sufficient for the purposes of this dissertation. Because of this form of the RBF, the interpolation nodes will often be referred to as centres of the RBF.

Recall that because of the Mairhuber example (Figure 1.1), the space the interpolant is to be chosen from should be dependent on the locations of the nodes. This fundamental property is directly built into RBFs. The question that naturally arises now is which choices of ϕ, with what degree of polynomial, lead to an invertible interpolation system.

Consider solving the Interpolation Problem 1.1 via an RBF with basic function ϕ augmented with a polynomial of degree $m-1$. That is, consider finding a function s of the

Name		Formula	SCPD
linear		$\phi(r) = -r$	$\mathrm{SCPD}_1(\mathbb{R}^d)$
cubic		$\phi(r) = r^3$	$\mathrm{SCPD}_2(\mathbb{R}^d)$
thin-plate spline		$\phi(r) = r^2 \ln r$	$\mathrm{SCPD}_2(\mathbb{R}^d)$
polyharmonic	^{a,b}	$\phi(r) = \begin{cases} (-1)^m r^{2m} \ln r, & \text{or} \\ (-1)^{m+1} r^{2m+1} \end{cases}$	$\mathrm{SCPD}_{m+1}(\mathbb{R}^d)$
multiquadric		$\phi(r) = -\sqrt{r^2+c^2}, \quad c>0$	$\mathrm{SCPD}_1(\mathbb{R}^d)$
inverse multiquadric		$\phi(r) = \left(r^2+c^2\right)^{-1/2}, \quad c>0$	$\mathrm{SCPD}_0(\mathbb{R}^d)$
generalised multiquadric	^{c}	$\phi(r) = (-1)^k \left(r^2+c^2\right)^{m/2}, \quad c>0$ and m is odd	(see note ^{c})
Gaussian		$\phi(r) = \exp\left(-cr^2\right), \quad c>0$	$\mathrm{SCPD}_0(\mathbb{R}^d)$
compactly supported	^{d}	$\phi(r) = (1-r)_+^{\lfloor d/2\rfloor+1}$	$\mathrm{SCPD}_0(\mathbb{R}^d)$

[a] A polyharmonic spline is a function $s : \mathbb{R}^d \to \mathbb{R}$ of the form (1.18) where

$$\phi(r) = \begin{cases} r^{2m-d} \ln r, & \text{if } 2m \geq d \text{ and } d \text{ is even,} \\ r^{2m-d}, & \text{otherwise,} \end{cases}$$

p is a polynomial of degree $m-1$ or less and the conditions (1.19b) are imposed on the coefficients. For such a function, $\Delta^m s = 0$ (a.e.).

[b] The positive definiteness is stated for $m \geq 0$. In the case that $m < 0$, the functions given by $\phi(r) = r^{2m} \ln r$ and $\phi(r) = r^{2m+1}$ are not defined at zero and hence cannot be (S)CPD.

[c] If $m > 0$ then $k = (m-1)/2$ and ϕ is $\mathrm{SCPD}_{(m+1)/2}(\mathbb{R}^d)$. If $m < 0$ then $k = 0$ and ϕ is SCPD.

[d] There are a number of piecewise polynomials that are $2k$ times continuously differentiable, compactly supported and suitable for use as RBF basic functions (see Wendland [69, 70] and Wu [72]). Those of minimal degree are given by

$$\phi(r) = \phi_{d,k}(r) = \int_{t_k=r}^{\infty} \int_{t_{k-1}=t_k}^{\infty} \cdots \int_{t_1=t_2}^{\infty} t_k \cdots t_2 t_1 (1-t_1)_+^{\lfloor d/2\rfloor+k+1} dt_1 dt_2 \cdots dt_k$$

and are constructed so that $\phi_{d,k} \in \mathrm{SCPD}_0(\mathbb{R}^d) \cap C^{2k}(\mathbb{R}^d)$.

Table 1.2: Some popular choices of basic functions for RBFs. The third column indicates the order of conditional positive definiteness (see Definition 1.14 to come).

form (1.18) that satisfies

$$s(x_i) = f_i, \qquad i = 1, \ldots, N, \tag{1.19a}$$

$$\sum_{i=1}^{N} \lambda_i q(x_i) = 0, \qquad \text{for all } q \in \pi_{m-1}(\mathbb{R}^d). \tag{1.19b}$$

In the case of the polyharmonic splines that come from solving Problem 1.3, the justification of the orthogonality conditions (1.19b) is clear. The orthogonality conditions can also have the effect of controlling the growth of the RBF away from the centres. This can often be seen via a far field expansion. Beatson, Cherrie & Mouat [7] use this technique to control the growth of approximate cardinal functions. Another interpretation is that the orthogonality conditions take away the extra degrees of freedom (dimensionality) added by the polynomial part.

The approach in the next few paragraphs is based on that of Sibson & Stone [62]. For the given nodes $X = \{x_i\}_{i=1}^{N}$, define the $N \times N$ matrix $\mathbf{\Phi}$ by $\mathbf{\Phi}_{i,j} = \phi(|x_i - x_j|)$. Choose a basis $\{p_j\}_{j=1}^{\ell_m}$ for the polynomial space π_{m-1} and define the $N \times \ell_m$ matrix P by $\mathsf{P}_{i,j} = p_j(x_i)$. Finding the interpolating RBF corresponds to finding the two vectors of coefficients: the coefficients $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)^\mathsf{T}$ of the radial part and the coefficients $\boldsymbol{c}$ of the polynomial part with respect to the basis $\{p_j\}_{j=1}^{\ell_m}$. Equations (1.19) may now be written in the matrix form

$$\begin{pmatrix} \mathbf{\Phi} & \mathsf{P} \\ \mathsf{P}^\mathsf{T} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ \boldsymbol{c} \end{pmatrix} = \begin{pmatrix} \boldsymbol{f} \\ 0 \end{pmatrix}, \tag{1.20}$$

where $\boldsymbol{f} = (f_1, \ldots, f_N)^\mathsf{T}$ is the vector of given values to interpolate to. If there is no polynomial part, only the smaller system $\mathbf{\Phi}\boldsymbol{\lambda} = \boldsymbol{f}$ need be solved.

Let Q be any (real valued) $N \times (N - \ell_m)$ matrix such that the columns of Q span the null space of P^T. Note that $\{p_j\}_{j=1}^{\ell_m}$ is not necessarily a Lagrange basis for $\pi_{m-1}(\mathbb{R}^d)$ and so P and Q may be different from those in the proof of Theorem 1.10, although, their roles are the same. The coefficients $\boldsymbol{\lambda} \in \mathbb{R}^N$ satisfy the orthogonality conditions of Equation (1.19b) if and only if there exists a $\boldsymbol{\mu} \in \mathbb{R}^{N-\ell_m}$ such that $\boldsymbol{\lambda} = \mathsf{Q}\boldsymbol{\mu}$. Equation (1.20) may now be rewritten as

$$\mathbf{\Phi}\mathsf{Q}\boldsymbol{\mu} + \mathsf{P}\boldsymbol{c} = \boldsymbol{f}.$$

Left multiplication of this expression by Q^T gives the reduced system

$$\mathsf{Q}^\mathsf{T}\mathbf{\Phi}\mathsf{Q}\boldsymbol{\mu} = \mathsf{Q}^\mathsf{T}\boldsymbol{f}, \tag{1.21}$$

since $Q^T P = 0$. If the matrix $Q^T \Phi Q$ is a definite matrix, either positive or negative, then Equation (1.21) always has a unique solution $\boldsymbol{\mu}$. Setting $\boldsymbol{\lambda} = Q\boldsymbol{\mu}$ and rewriting Equation (1.21) as

$$0 = Q^T (\boldsymbol{f} - \Phi\boldsymbol{\lambda}),$$

it follows that $\boldsymbol{f} - \Phi\boldsymbol{\lambda}$ is in the column space of P and hence there is a unique solution $\boldsymbol{c}$ to the system

$$P\boldsymbol{c} = \boldsymbol{f} - \Phi\boldsymbol{\lambda}.$$

There thus exists a unique solution $\{\boldsymbol{\lambda}, \boldsymbol{c}\}$ to Equation (1.20) if the matrix $Q^T \Phi Q$ is a definite matrix, *i.e.*, there is a unique solution if

$$\boldsymbol{\lambda}^T \Phi \boldsymbol{\lambda} > 0 \qquad \text{for all } \boldsymbol{\lambda} \in \mathbb{R}^N \backslash \{0\} \text{ such that } P^T \boldsymbol{\lambda} = 0 .$$

Equivalently '< 0' could be used rather than '> 0.' Definition 1.14 implies that if ϕ is strictly conditionally positive definite of order m on $\mathbb{R}^d$ then Equation (1.20) has a unique solution for any choice of the node set $X = \{x_i\}_{i=1}^N \subset \mathbb{R}^d$ that is π_{m-1}-unisolvent.

Definition 1.14. A continuous function ϕ, defined on $[0, \infty)$, is said to be *conditionally positive definite* (CPD) of order m on $\mathbb{R}^d$ if for any distinct points $x_1, \ldots, x_N \in \mathbb{R}^d$ and for any $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)^T \in \mathbb{R}^N$ such that

$$\sum_{i=1}^N \lambda_i p(x_i) = 0, \qquad \text{for all } p \in \pi_{m-1}(\mathbb{R}^d), \tag{1.22}$$

the quadratic form

$$\sum_{i,j=1}^N \lambda_i \lambda_j \phi(|x_i - x_j|) \tag{1.23}$$

is non-negative. If the given quadratic form is positive for all non-zero choices of $\boldsymbol{\lambda}$ that satisfy Equation (1.22) then ϕ is said to be *strictly* conditionally positive definite (SCPD) of order m on $\mathbb{R}^d$.

Note that this definition may disagree with that of some authors. For example, Micchelli [49] uses the quadratic form $\sum_{i,j=1}^n \lambda_i \lambda_j \phi(|x_i - x_j|^2)$ instead of Equation (1.23). Thus if $\widetilde{\phi}$ is (S)CPD by Micchelli's definition, then $\phi(\cdot) = \widetilde{\phi}(\cdot^2)$ is (S)CPD by the above definition.

The class of conditionally positive definite functions of order m is denoted by $\mathrm{CPD}_m(\mathbb{R}^d)$. Since $\pi_{m-1}(\mathbb{R}^d) \subset \pi_m(\mathbb{R}^d)$, it follows that $\mathrm{CPD}_{m+1}(\mathbb{R}^d) \subseteq \mathrm{CPD}_m(\mathbb{R}^d)$. Similarly for the strictly conditionally positive definite functions $\mathrm{SCPD}_m(\mathbb{R}^d)$.

Example 1.15. If ϕ and ψ are both members of $\mathrm{SCPD}_m(\mathbb{R}^d)$ then so is $\phi+\psi$. Furthermore, for any scalar $a>0$, the function $a\phi$ is also in $\mathrm{SCPD}_m(\mathbb{R}^d)$.

Example 1.16. The function $\phi(r) = r^2$ is CPD of order 1 on $\mathbb{R}^d$ for any d. However, it is *not* SCPD. This means there is a data set $\{x_i, f_i\}_{i=1}^N$ such that the Interpolation Problem 1.1 cannot be solved by an RBF with this basic function. This should come as no surprise as $\mathrm{span}\{|\cdot - x_i|^2 : x_i \in X\}$ is at most just the space of quadratics on $\mathbb{R}^d$ and hence has dimension at most $(d+2)(d+1)/2$. Equivalent statements may be made for other choices of $\phi(r) = r^{2m}$, where m is a non-negative integer.

Example 1.17. For any $c > 0$, the Gaussian $\phi(r) = e^{-cr^2}$ is strictly positive definite. This is an immediate consequence of Bochner's Theorem (see [19, p. 82] or [63, p. 413]).

The following example of a strictly conditionally positive definite function will prove to be very important in showing that a number of other functions are also strictly conditionally positive definite.

Lemma 1.18. (Powell [54, Theorem 3.3]) *Let w be a non-negative function such that*

$$\phi(r) = \int_0^\infty e^{-r^2 t} w(t)dt \tag{1.24}$$

is well defined for all $r \geq 0$ and

$$\int_0^\infty w(t)dt > 0. \tag{1.25}$$

Then ϕ is strictly conditionally positive definite of order 0 on $\mathbb{R}^d$ for all d.

Proof. Since $\phi(0) = \int_0^\infty w(t)dt$ is well defined, w is absolutely integrable on $(0,\infty)$. Direct calculation of the quadratic form (1.23) gives

$$\begin{aligned}\sum_{i,j=1}^N \lambda_i\lambda_j\phi(|x_i - x_j|) &= \sum_{i,j=1}^N \lambda_i\lambda_j \int_0^\infty \exp(-|x_i - x_j|^2 t)w(t)dt \\ &= \int_0^\infty \left(\sum_{i,j=1}^N \lambda_i\lambda_j \exp(-|x_i - x_j|^2 t)\right) w(t)dt.\end{aligned}$$

The expression in the parenthesis is positive for all positive values of t by Example 1.17 and for $t = 0$ is equal to $\left(\sum_{i=1}^N \lambda_i\right)^2$, and is therefore non-negative. Since w is non-negative and not identically zero, it follows that the integral, and thus the quadratic form (1.23), is positive. ∎

Example 1.19. Since

$$\left(r^2 + c^2\right)^{-1/2} = \int_0^\infty e^{-tr^2} \underbrace{\pi^{-1/2} t^{-1/2} e^{-tc^2}}_{w(t)} dt$$

(Powell [54, pp. 121]), Lemma 1.18 implies that the inverse multiquadric is strictly positive definite.

Definition 1.20. A function f is said to be *completely monotonic* on $(0, \infty)$ provided that it is in $C^\infty(0, \infty)$ and

$$(-1)^\ell f^{(\ell)}(x) \geq 0, \qquad x \in (0, \infty), \quad \ell = 0, 1, 2, \ldots.$$

Example 1.21. If p is a polynomial and p is completely monotonic on $(0, \infty)$, then p is constant.

This is easily shown by induction on the degree on p. For the induction basis, consider the case when p is linear. Specifically, assume that $p(x) = ax + b$. Since $p'(x) \leq 0$ on $(0, \infty)$ it follows that $a \leq 0$. Now, $p(x) \geq 0$ implies that $ax \geq -b$ for all $x \geq 0$, *i.e.*, $a = 0$. Thus, if p is both a linear function and completely monotonic, then p is constant.

For the induction step, notice that if f is completely monotonic then so is $-f'$. Hence, the inductive hypothesis and the complete monotonicity of a polynomial p imply that the derivative p' is a constant and thus p is linear. However, if p is linear and completely monotonic, then, by the induction basis, p is constant.

Completely monotonic functions may be characterised by the following theorem of Bernstein [71, Theorem 12b].

Theorem 1.22. *A necessary and sufficient condition that f be completely monotonic on $(0, \infty)$ is that*

$$f(t) = \int_0^\infty e^{-t\sigma} d\mu(\sigma)$$

where μ is non-decreasing and the integral converges for all $t \in (0, \infty)$.

Micchelli [49] proved the following result.

Theorem 1.23. *Let $\widetilde{\phi} \in C^m[0, \infty)$ be such that $(-1)^m \widetilde{\phi}^{(m)}$ is completely monotonic on $(0, \infty)$ but not constant. Then $\phi(\cdot) = \widetilde{\phi}(\cdot^2)$ is strictly conditionally positive definite of order m on $\mathbb{R}^d$ for all d.*

Since this result is so important to RBF theory, we give an expanded version of Micchelli's argument below. The proof of this will require Taylor's theorem with integral remainder.

Theorem 1.24. *Let $f \in C^m[a, b]$ and let $t_0 \in [a, b]$. Then for all $a \leq t \leq b$*

$$f(t) = \sum_{\ell=0}^{m-1} \frac{f^{(\ell)}(t_0)}{\ell!}(t - t_0)^\ell + \frac{1}{(m-1)!} \int_{t_0}^t f^{(m)}(\tau)(t - \tau)^{m-1} d\tau.$$

Lemma 1.25. (Micchelli [49]). *Let d and N be a positive integers and $x_1, \ldots, x_N \in \mathbb{R}^d$ be any distinct points. Furthermore, let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)^\mathsf{T} \in \mathbb{R}^N$ satisfy the orthogonality conditions* (1.22)*. Then*

$$\sum_{i,j=1}^{N} \lambda_i \lambda_j |x_i - x_j|^{2\ell} = 0, \qquad 0 \le \ell \le m-1.$$

Proof. Using the binomial theorem twice,

$$\begin{aligned} |x_i - x_j|^{2\ell} &= \sum_{\alpha=0}^{\ell} \sum_{\beta=0}^{\alpha} \binom{\ell}{\alpha} \binom{\alpha}{\beta} (2\langle x_i, x_j \rangle)^{\ell-\alpha} |x_i|^{2\beta} |x_j|^{2(\alpha-\beta)} \\ &= \sum_{\alpha=0}^{\ell} \sum_{\beta=0}^{\alpha} p_{\alpha,\beta}(x_i, x_j). \end{aligned} \tag{1.26}$$

Since $p_{\alpha,\beta}(\cdot, x_j) \in \pi_{\ell-(\alpha-2\beta)}(\mathbb{R}^d)$ the orthogonality conditions on $\boldsymbol{\lambda}$ imply

$$\sum_{i=1}^{N} \lambda_i p_{\alpha,\beta}(x_i, x_j) = 0, \qquad \text{for } \alpha - 2\beta \ge 0 \text{ and all } x_j \in \mathbb{R}^d. \tag{1.27}$$

Similarly, since $p_{\alpha,\beta}(x_i, \cdot) \in \pi_{\ell+(\alpha-2\beta)}(\mathbb{R}^d)$,

$$\sum_{j=1}^{N} \lambda_j p_{\alpha,\beta}(x_i, x_j) = 0, \qquad \text{for } \alpha - 2\beta \le 0 \text{ and all } x_i \in \mathbb{R}^d. \tag{1.28}$$

Together, Equations (1.27) and (1.28) imply that

$$\sum_{i,j=1}^{N} \lambda_i \lambda_j p_{\alpha,\beta}(x_i, x_j) = 0, \qquad 0 \le \beta \le \alpha, \quad 0 \le \alpha \le \ell.$$

Thus, multiplying both sides of Equation (1.26) by $\lambda_i \lambda_j$ and summing over i and j gives

$$\sum_{i,j=1}^{N} \lambda_i \lambda_j |x_i - x_j|^{2\ell} = \sum_{\alpha=0}^{\ell} \sum_{\beta=0}^{\alpha} \sum_{i,j=1}^{N} \lambda_i \lambda_j p_{\alpha,\beta}(x_i, x_j) = 0.$$

∎

Proof of Theorem 1.23. By the complete monotonicity of $(-1)^m \widetilde{\phi}^{(m)}$ there is a Borel measure $d\mu$ such that

$$(-1)^m \widetilde{\phi}^{(m)}(t) = \int_0^\infty e^{-t\sigma} d\mu(\sigma). \tag{1.29}$$

For $\epsilon > 0$, define $\widetilde{\phi}_\epsilon = \widetilde{\phi}(\cdot + \epsilon)$. Then $\widetilde{\phi}_\epsilon \in C^m[0,\infty)$ and Taylor's Theorem 1.24 may be applied to $\widetilde{\phi}_\epsilon$ with $t_0 = 0$. Thus

$$\widetilde{\phi}_\epsilon(t) - \sum_{\ell=0}^{m-1} \frac{\widetilde{\phi}_\epsilon^{(\ell)}(0)}{\ell!} t^\ell = \frac{1}{(m-1)!} \int_0^t \widetilde{\phi}_\epsilon^{(m)}(\tau)(t-\tau)^{m-1} d\tau.$$

Substituting Equation (1.29) into this expression and interchanging the order of integration gives

$$\widetilde{\phi}_\epsilon(t) - \sum_{\ell=0}^{m-1} \frac{\widetilde{\phi}_\epsilon^{(\ell)}(0)}{\ell!} t^\ell = \frac{(-1)^m}{(m-1)!} \int_0^\infty e^{-\epsilon\sigma} \int_0^t e^{-\tau\sigma}(t-\tau)^{m-1} d\tau d\mu(\sigma). \tag{1.30}$$

Using integration by parts a number of times, the inner integral is easily evaluated:

$$\begin{aligned} \frac{(-1)^m}{(m-1)!} \int_0^t e^{-\tau\sigma}(t-\tau)^{m-1} d\tau &= \frac{(-1)^m}{(m-1)!} \frac{t^{m-1}}{\sigma} + \frac{(-1)^{m-1}}{(m-2)!} \frac{1}{\sigma} \int_0^t e^{-\tau\sigma}(t-\tau)^{m-2} d\tau \\ &= \sum_{\ell=1}^{m-1} \frac{(-1)^{m+1-\ell}}{(m-\ell)!} \frac{t^{m-\ell}}{\sigma^\ell} + \frac{1}{\sigma^{m-1}} \int_0^t e^{-\tau\sigma} d\tau \\ &= \sum_{\ell=1}^{m-1} \frac{(-1)^{m+1-\ell}}{(m-\ell)!} \frac{t^{m-\ell}}{\sigma^\ell} + \frac{e^{-t\sigma}}{\sigma^m} - \frac{1}{\sigma^m} \\ &= \frac{1}{\sigma^m} \left\{ e^{-t\sigma} - \sum_{\ell=0}^{m-1} \frac{(-t\sigma)^\ell}{\ell!} \right\}. \end{aligned}$$

Substituting this expression back into Equation (1.30) gives

$$\widetilde{\phi}_\epsilon(t) - \sum_{\ell=0}^{m-1} \frac{\widetilde{\phi}_\epsilon^{(\ell)}(0)}{\ell!} t^\ell = \int_0^\infty \frac{e^{-\epsilon\sigma}}{\sigma^m} \left\{ e^{-t\sigma} - \sum_{\ell=0}^{m-1} \frac{(-t\sigma)^\ell}{\ell!} \right\} d\mu(\sigma). \tag{1.31}$$

Let d and N be a positive integers and choose any distinct points $x_1, \ldots, x_N \in \mathbb{R}^d$. Furthermore, choose any $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)^\top \in \mathbb{R}^N$ that satisfies the orthogonality condition (1.22). In Equation (1.31), set $t = |x_i - x_j|^2$, multiply both sides by $\lambda_i\lambda_j$ and sum over i and j, *i.e.*, consider

$$\begin{aligned} &\sum_{i,j=1}^N \lambda_i\lambda_j \widetilde{\phi}_\epsilon(|x_i - x_j|^2) - \sum_{\ell=0}^{m-1} \frac{\widetilde{\phi}_\epsilon^{(\ell)}(0)}{\ell!} \sum_{i,j=1}^N \lambda_i\lambda_j |x_i - x_j|^{2\ell} \\ &= \int_0^\infty \frac{e^{-\epsilon\sigma}}{\sigma^m} \left\{ \sum_{i,j=1}^N \lambda_i\lambda_j e^{-|x_i-x_j|^2\sigma} - \sum_{\ell=0}^{m-1} \frac{(-\sigma)^\ell}{\ell!} \sum_{i,j=1}^N \lambda_i\lambda_j |x_i - x_j|^{2\ell} \right\} d\mu(\sigma). \end{aligned}$$

Lemma 1.25 may be used to reduce this expression to

$$\sum_{i,j=1}^N \lambda_i\lambda_j \widetilde{\phi}_\epsilon(|x_i - x_j|^2) = \int_0^\infty \frac{e^{-\epsilon\sigma}}{\sigma^m} \sum_{i,j=1}^N \lambda_i\lambda_j e^{-\sigma|x_i-x_j|^2} d\mu(\sigma). \tag{1.32}$$

By considering the Taylor expansion of the exponential function, it is easily seen that the orthogonality of the coefficients $\boldsymbol{\lambda}$ to polynomials implies that

$$\sum_{i,j=1}^N \lambda_i\lambda_j e^{-\sigma|x_i-x_j|^2} = \mathcal{O}(\sigma^m), \qquad \text{as } \sigma \to 0,$$

and so the integral in Equation (1.32) is well defined.

Since $\widetilde{\phi}^{(m)}$ is non-constant, the measure μ must have weight away from zero. That is, there exists an $a > 0$ such that

$$\int_a^{2a} d\mu(\sigma) > 0$$

Let ϵ be so small that $\epsilon < \log(2)/2a$. Then $\exp(-\epsilon\sigma) > 1/2$ for all $\sigma \in [a, 2a]$. Furthermore, by Example 1.17, the quadratic form under the integral in Equation (1.32) is strictly greater than zero on the interval $[a, 2a]$. By the continuity of that quadratic as a function of σ, it is bounded away from 0 on $[a, 2a]$. Thus, for all ϵ in the range $0 < \epsilon < \log(2)/2a$,

$$\int_0^\infty \frac{e^{-\epsilon\sigma}}{\sigma^m} \sum_{i,j=1}^N \lambda_i\lambda_j e^{-\sigma|x_i-x_j|^2} d\mu(\sigma) \geq \frac{1}{2}\int_a^{2a} \sigma^{-m} \sum_{i,j=1}^N \lambda_i\lambda_j e^{-\sigma|x_i-x_j|^2} d\mu(\sigma) = \delta > 0.$$

Substituting this expression into Equation (1.32) gives

$$\sum_{i,j=1}^N \lambda_i\lambda_j\widetilde{\phi}_\epsilon(|x_i - x_j|^2) \geq \delta > 0, \qquad 0 < \epsilon < \frac{\log(2)}{2a}.$$

Since $\widetilde{\phi}$ is continuous at 0, taking the limit as $\epsilon \to 0^+$ and substituting $\phi = \widetilde{\phi}(\cdot^2)$ gives

$$\sum_{i,j=1}^N \lambda_i\lambda_j\phi(|x_i - x_j|) > 0.$$

∎

Example 1.26. The function given by $\phi(r) = r$ is $\mathrm{SCPD}_1(\mathbb{R}^d)$ for all positive integers d. The derivatives of $\widetilde{\phi} = \phi(\sqrt{\cdot})$ are given by

$$\widetilde{\phi}^{(k)}(r) = \begin{cases} \frac{1}{2}r^{-1/2}, & k = 1, \\ (-1)^{k+1}\frac{(2k-3)!!}{2^k}r^{(2k-1)/2}, & k = 2, 3, \ldots, \end{cases}$$

where $(2k-3)!! = (2k-3)(2k-5)\ldots(3)(1)$. Hence it follows that $\widetilde{\phi}'$ is completely monotonic on $(0, \infty)$ and thus, by Theorem 1.23, ϕ is $\mathrm{SCPD}_1(\mathbb{R}^d)$ for all d.

The following theorem is the converse to Theorem 1.23.

Theorem 1.27. (Guo, Hu and Sun [30, Theorem 2.1]) *Let $\widetilde{\phi}$ be a continuous function on $[0, \infty)$, and let $\phi(\cdot) = \widetilde{\phi}(\cdot^2)$. The following two statements are equivalent.*

(a) *The function ϕ is conditionally positive definite of order m on $\mathbb{R}^d$ for all positive integers d.*

(b) *The derivative $\widetilde{\phi}^{(j)}(t)$ exists for all positive integers j and for all $t \in (0, \infty)$, and furthermore, $(-1)^m\widetilde{\phi}^{(m)}$ is completely monotonic on $(0, \infty)$.*

The non-unisolvent case

Theorem 1.23 ensures invertibility of the RBF interpolation system (1.20) in the case that the node set X is unisolvent with respect to $\pi_{m-1}(\mathbb{R}^d)$. However, when techniques such as the domain decomposition method of Beatson, Light & Billings [11] are used, it is not uncommon to find that the node set is not unisolvent. We will now briefly describe one way in which this potentially troublesome case may be overcome.

Formally, given any subset $X \subset \mathbb{R}^d$ we define $\pi_{m-1}(X)$ as a space of equivalence classes of polynomials from $\pi_{m-1}(\mathbb{R}^d)$, with the equivalence relation $\equiv$ being equality on X. Then a basis for $\pi_{m-1}(X)$ is a spanning set of representatives linearly independent over X.

Given X not unisolvent for $\pi_{m-1}(\mathbb{R}^d)$ choose a basis $\{q_j\}_{j=1}^k$ for $\pi_{m-1}(X)$. Define $\mathcal{P} : \pi_{m-1}(\mathbb{R}^d) \to \pi_{m-1}(X)$ by

$$\mathcal{P}q \in \operatorname{span}\{q_j\}_{j=1}^k \qquad \text{and} \qquad \mathcal{P}q(x) = q(x) \quad \text{for all } x \in X.$$

The operator $\mathcal{P}$ is a projection from $\pi_{m-1}(\mathbb{R}^d)$ onto $Q_k = \operatorname{span}\{q_j\}_{j=1}^k$. Hence $(I - \mathcal{P})$ is a projection from $\pi_{m-1}(\mathbb{R}^d)$ onto the space of polynomials that are identically zero on X. This space has dimension $\ell_m - k$. Choose $\{q_j\}_{j=k+1}^{\ell_m}$ a basis for $(I - \mathcal{P})\pi_{m-1}(\mathbb{R}^d)$. Now

$$\sum_{i=1}^{N} \lambda_i q(x_i) = 0 \qquad \text{for all } q \in \pi_{m-1}(\mathbb{R}^d)$$

if and only if

$$\sum_{i=1}^{k} \lambda_i q(x_i) = 0 \qquad \text{for all } q \in Q_k. \tag{1.33}$$

Hence the orthogonality conditions (1.19b) may be replaced by new conditions (1.33) and the correspondingly modified matrix P in Equation (1.20) is full rank even in the non-unisolvent case.

1.2 Fast evaluation via the Fast Multipole Method

This section gives an introduction to the hierarchical and fast multipole methods that are the basis of fast evaluators. The specifics of these methods are the focus of Chapters 2 and 3 of this thesis.

As was shown in the previous section, radial basis functions provide a good way to solve the interpolation problem (Problem 1.1). The main reason for this is that existence and uniqueness of an interpolant is guaranteed under mild conditions on the locations of the interpolation nodes. For example, unisolvency of the interpolation nodes with respect to low degree polynomials is sufficient. Furthermore, if desired, an RBF can be chosen to satisfy certain energy minimization conditions (such as those of Problem 1.3).

However, there are clear computational difficulties if an RBF is to be used to solve the interpolation problems of Chapter 4 where there may be in excess of 70,000 interpolation nodes. Recall that if the RBF of Equation (1.18) is to solve an interpolation problem the coefficients λ_i are found by solving the linear system of Equation (1.20). If N is large, *i.e.,* in the tens or hundreds of thousands, then direct solution of this system is out of the question. Moreover, if an iterative method, such as GMRES [7, 57], is used then a matrix-vector product is required at each iteration step. The particular product that is required is equivalent to evaluation of an RBF at all of its centres. In view of Equation (1.18), this process would seem to require $\mathcal{O}(N^2)$ flops. Using the Fast Multipole Method (FMM) the cost of evaluating an RBF with N centres can be reduced to $\mathcal{O}(N \log N)$ flops.

Once the coefficients of the RBF have been determined, it is generally required to evaluate it at some large number of points. For example, once an interpolating surface has been found, it may need to be evaluated enough times to produce a plot of the surface. Once again, the naïve estimate for a single evaluation of an N centre RBF is $\mathcal{O}(N)$ flops. With the FMM the cost of one extra evaluation can be reduced to $\mathcal{O}(1)$ flops with an initial setup of $\mathcal{O}(N \log N)$ flops.

The FMM was developed by Greengard & Rokhlin [29] for the evaluation of potentials in two and three dimensions. Since the potential due to a single source is rotation invariant, the potential function they wished to evaluate is in fact an RBF. Furthermore, their method extends to more general cases of RBFs. For instance, the FMM may be applied to polyharmonic RBFs in two [10, 12], three [13, 25], or four dimensions [8], and to the generalised multiquadric in arbitrary dimensions [20].

1.2.1 An hierarchical evaluator

When computations are performed on real data, infinite precision is neither expected nor required. What is expected is that a given level of precision is achieved. The FMM exploits this fact; once it is decided what level of precision is required the FMM constructs an approximation to the RBF that can be evaluated comparatively quickly and that gives a result within this precision.

The first step in a hierarchical evaluator or FMM is to construct an approximation to a single basic function. What is desired is an approximation such that the sum of many basic functions can be evaluated at cost that depends only on the accuracy of the approximation not the number of basic functions in the sum. This can be achieved by a truncated *far field* expansion, which is an expansion that is valid far away from the centre of expansion. It is a Laurent-like expansion; it is valid outside a certain ball.

The FMM will be demonstrated using the ordinary multiquadric RBF of one variable as an example. From Lemma 3.6 and Example 3.7, the far field expansion for the multiquadric basic function is given by

$$\Phi(x-t) = \sqrt{(x-t)^2+c^2} = \operatorname{sign}(x)\sum_{\ell=0}^{\infty} P_\ell(t^2+c^2,-2t,1)/x^{\ell-1}, \qquad |x| > \sqrt{t^2+c^2},$$

where the polynomials $P_\ell = P_\ell^{(1)}$ are given by Equation (3.4). The approximation to $\Phi(\cdot - t)$ will be constructed by truncating the given series, dropping all terms that are $\mathcal{O}(|x|^{-p-1})$ as $x \to \infty$, *i.e.*,

$$g_p(x) = \operatorname{sign}(x)\sum_{\ell=0}^{p+1} P_\ell(t^2+c^2,-2t,1)/x^{\ell-1}$$

is the truncated far field approximation to $\Phi(\cdot - t)$. Lemma 3.6 provides a bound on the error due to approximating $\Phi(\cdot - t)$ with g_p. Specifically,

$$\left|\Phi(x-t) - g_p(x)\right| \le 2\sqrt{t^2+c^2}\left(\frac{\sqrt{t^2+c^2}}{|x|}\right)^{p+1}\frac{|x|}{|x|-\sqrt{t^2+c^2}}, \qquad |x| > \sqrt{t^2+c^2}. \quad (1.34)$$

By choosing p sufficiently large any level of accuracy may be obtained for the approximation g_p.

Assume that the *source point* t is inside the unit sphere and that the *evaluation point* is far away from the origin, $|x| > 3$, say. Using $c = 0.1$ for the multiquadric constant, if the condition

$$\left|\Phi(x-t) - g_p(x)\right| < 10^{-4},$$

is to hold then it is required that $p \ge 15$. Clearly it will be more expensive to evaluate g_{15} than $\Phi(\cdot - t)$. However, to compute the sum

$$\sum_{i=1}^{N} \lambda_i \Phi(x - x_i)$$

for large N it is more efficient to compute the approximation

$$\operatorname{sign}(x)\sum_{\ell=0}^{p+1} C_\ell / x^{\ell-1}, \qquad (1.35)$$

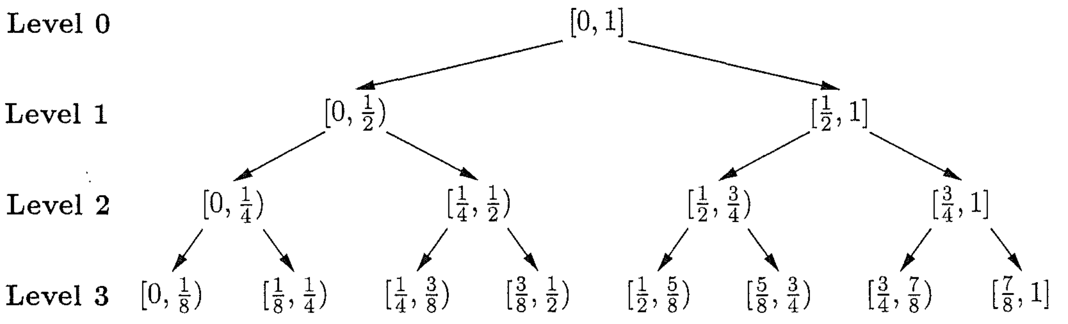


Figure 1.7: Binary tree structure induced by a uniform subdivision of the unit interval.

provided that the coefficients

$$C_\ell = \sum_{i=1}^{N} \lambda_i P_\ell(x_i^2 + c^2, -2x_i, 1), \qquad \ell = 0, \ldots, p+1,$$

have been precalculated. Note that the cost of evaluating (1.35) does not increase when the number of source points increases; it only increases when the desired accuracy increases.

Unfortunately, in most cases the entire RBF cannot be replaced with an approximation such as (1.35). This approximation is only valid when the evaluation point is sufficiently far from the source points. However, by grouping the source points into clusters and using a careful combination of the direct and approximate forms, the desired reduction in the cost of evaluation can be achieved. The clustering of source points is easily and suitably achieved with an hierarchical division of space.

In an hierarchical division of space, each region of space, referred to as a *panel*, is subdivided into child panels. The union of the child panels should equal the parent panel. The largest and coarsest panel should be the whole domain and have no parents. As with the accuracy of computation, an infinitely deep tree of panels is neither required, nor possible. Thus some suitable stoping criteria is used to decide when a panel is too small to be divided into child panels. "Smallness" may be based on the actual size of the panel or some other criteria such as the number of source points in a panel. Figure 1.7 shows a simple hierarchical division of the unit interval. In this example, $[0, \frac{1}{4})$ has two children, $[0, \frac{1}{8})$ and $[\frac{1}{8}, \frac{1}{4})$, and its parent is $[0, \frac{1}{2})$.

Associated with each panel in the hierarchical division is the part of the RBF due to centres inside that panel and also an approximation to that part of the RBF. That is, associated with panel T will be the RBF

$$s_T(x) = \sum_{i:x_i \in T} \lambda_i \phi(|x - x_i|),$$

and an approximation to s_T, denoted r_T. For a panel T centred on the origin, the approximation to s_T is given by

$$r_T(x) = \operatorname{sign}(x)\sum_{\ell=0}^{p+1} C_{T,\ell}/x^{\ell-1}, \qquad C_{T,\ell} = \sum_{i:x_i\in T} \lambda_i P_\ell(x_i^2+c^2, -2x_i, 1), \tag{1.36}$$

where p has been chosen so that $|s_T(x) - r_T(x)|$ is sufficiently small outside some ball that contains T and is centred at the origin. If T is centred, not at the origin but at some other point, u say, then x and x_i may be replaced by $x-u$ and x_i-u respectively in Equation (1.36) and the region of validity for the approximation, *i.e.*, where $|s_T(x)-r_T(x)|$ is sufficiently small, is now the exterior of some ball centred on u. Note that the choice of p is dependent both on the required precision and on the chosen radius (with respect to the size of T) of the ball outside of which the approximation is valid.

In an implementation of this hierarchical evaluator, the following will be associated with each panel:

- part of the RBF, $s_T(x) = \sum_{i:x_i\in T} \lambda_i \phi(|x-x_i|)$,
- a far field approximation to s_T, r_T,
- a ball B_T such that $|s_T(x) - r_T(x)|$ is sufficiently small for all $x \notin B_T$.

Note that the details of exactly how this information is stored is dependent upon the implementation and thus may take many different forms. This is especially true of the third point, which really refers to associating a region of validity with each panel, or more correctly with a given approximation r_T. For example, in an hierarchical division of space, it is computationally more efficient to have the boundaries of the various regions of validity lie along the boundaries of the panels. In this way checking to see if $r_T(x)$ is a good approximation to $s_T(x)$ can be done using the tree topology and without unnecessary distance calculations. With this type of strategy in place, if an evaluation panel T_1 does not intersect B_{T_2} for some source panel T_2 then $r_{T_2}(x)$ is a good approximation to $s_{T_2}(x)$ for all $x \in T_1$.

If $\mathcal{T}$ is a collection of pairwise disjoint panels which cover all of the centres in the RBF then

$$s(x) = p(x) + \sum_{T\in\mathcal{T}} s_T(x). \tag{1.37}$$

Furthermore, if $\{\mathcal{T}_1, \mathcal{T}_2\}$ is a partition of $\mathcal{T}$ then

$$p(x) + \sum_{T\in\mathcal{T}_1} s_T(x) + \sum_{T\in\mathcal{T}_2} r_T(x), \tag{1.38}$$

can approximate $s(x)$ to any desired precision for all x outside of $\bigcup_{T\in\mathcal{T}_2} B_T$.

We now return to the example problem of evaluating a ordinary multiquadric RBF of one variable with centres in the unit interval. Let the unit interval be divided up as in Figure 1.7. For each panel $T = [a, a+h)$ the ball B_T will be centred at $a+h/2$ and have radius $3h/2$. Let T_1 and T_2 be two panels at the same level. Furthermore, let the evaluation point x be in T_1 and the source points x_i be in T_2. Then, $r_{T_2}(x)$ is a sufficiently good approximation to $s_{T_2}(x)$ if T_1 and T_2 are separated by a third panel at the same level. If T_1 and T_2 are separated in this way, they are said to be *well separated.* The advantage of using this criteria is that, as mentioned above, it may be easily checked from the tree topology without reference to the actual layout of the panels and without distance computation.

For example, if $x \in [0, \frac{1}{8})$ then

$$s(x) \approx s_{[0,\frac{1}{8})}(x) + s_{[\frac{1}{8},\frac{1}{4})}(x) + r_{[\frac{1}{4},\frac{3}{8})}(x) + r_{[\frac{3}{8},\frac{1}{2})}(x) + r_{[\frac{1}{2},\frac{3}{4})}(x) + r_{[\frac{3}{4},1]}(x).$$

Obviously, $[0, \frac{1}{8})$ is not well separated from itself; neither is its neighbour $[\frac{1}{8}, \frac{1}{4})$. Thus, the approximates $r_{[0,\frac{1}{8})}(x)$ and $r_{[\frac{1}{8},\frac{1}{4})}(x)$ are not valid. However, both $[\frac{1}{4}, \frac{3}{8})$ and $[\frac{3}{8}, \frac{1}{2})$ are well separated from $[0, \frac{1}{8})$—the panel $[\frac{1}{8}, \frac{1}{4})$ is between each of these source panels and the evaluation panel $[0, \frac{1}{8})$—and hence the approximations $r_{[\frac{1}{4},\frac{3}{8})}(x)$ and $r_{[\frac{3}{8},\frac{1}{2})}(x)$ may be used. Finally, there are the panels that cover $[\frac{1}{2}, 1]$. This is where the computational savings have the most effect. Rather than using the panels at the finest level (Level 3), the panels at the next coarsest level may used since both $[\frac{1}{2}, \frac{3}{4})$ and $[\frac{3}{4}, 1]$ are separated from $[0, \frac{1}{4})$ by $[\frac{1}{4}, \frac{1}{2})$. Notice that the cost of evaluating r_T is independent of the panel size or the number of centres in T and depends only on the required level of precision. Thus it is cheaper to evaluate $r_{[\frac{1}{2},\frac{3}{4})}(x)$ than to evaluate $r_{[\frac{1}{2},\frac{5}{8})}(x) + r_{[\frac{5}{8},\frac{3}{4})}(x)$. Similarly, if $x \in [\frac{1}{2}, \frac{5}{8})$ then

$$s(x) \approx s_{[\frac{3}{8},\frac{1}{2})}(x) + s_{[\frac{1}{2},\frac{5}{8})}(x) + s_{[\frac{5}{8},\frac{3}{4})}(x) + r_{[\frac{1}{4},\frac{3}{8})}(x) + r_{[\frac{3}{4},\frac{7}{8})}(x) + r_{[\frac{7}{8},1]}(x) + r_{[0,\frac{1}{4})}(x).$$

A division of any space along the lines of Figure 1.7 is suitable when the centres of the RBF are uniformly distributed. A more general stopping criteria for the tree construction is to only split a panel if it contains more than some threshold number of centres. In the given case of uniformly distributed data, there would be $\mathcal{O}(\log N)$ levels in the tree. With the separation criteria used above for deciding between direct (s_T) and approximate (r_T) evaluation, there are at most three direct evaluations at the lowest level and at most three approximate evaluations on all levels, except Levels 0 and 1 where there is no evaluation. As the tree of panels has been constructed so that there are $\mathcal{O}(1)$ centres in each panel at the lowest level, the cost of evaluating s_T for T at this level is $\mathcal{O}(1)$. The cost of evaluating r_T

is $\mathcal{O}(p)$, independent of the level of T. Thus, the cost of evaluation at a single point by this hierarchical evaluator is $\mathcal{O}(p \log N)$.

This hierarchical evaluator may be written in pseudo-code as in Algorithm 1.1.

> **Input**: An evaluation point x and a panel T containing source points.
> **Output**: The value of that part of the RBF due to the source points in T evaluated at x.
> HIERARCHEVAL(x, T)
>
> **if** $x \notin B_T$
>   **return** $r_T(x)$
> **else if** T is at the lowest level
>   **return** $s_T(x)$
> **else**
>   **return** HIERARCHEVAL$(x, T_1) + \cdots +$ HIERARCHEVAL(x, T_n)
>   where $T_1, \ldots, T_n$ are the children of T.

Algorithm 1.1: Pseduo-code for a hierarchical evaluator.

A similar strategy may used for evaluating an RBF of two variables. An example of how a square might be divided up into a quad-tree of panels, each panel in turn being a square, is shown in Figure 1.8. Once again this is suitable for uniformly distributed data. The panels in Figure 1.8 are hatched to indicated how that part of the RBF corresponding to each panel would be evaluated in the case that the evaluation point is located as indicated. As with the one dimensional example, the idea of well separated panels has been used to determine whether approximation or direct evaluation is to be used in any particular case. In general, the cost of a single evaluation of an RBF of d variables can be reduced by this hierarchical evaluator from $\mathcal{O}(N)$ to $\mathcal{O}(p^d \log N)$ flops.

From examination of Equation (1.34), it can be seen that the error in approximating s_T by r_T is $\mathcal{O}\big((r/R)^{p+1}\big)$ as $p \to \infty$, where r is the radius of the panel T and R is the radius of the associated ball B_T. This is the case in general, not just in the specific example of the one-dimensional ordinary multiquadric (see, *e.g.,* Theorem 2.24, Theorem 2.32, or Theorem 3.9.) If a uniform division of space is to be used along with the given notion of well separateness, then there is a, perhaps surprising, dimensional limit to using the hierarchical

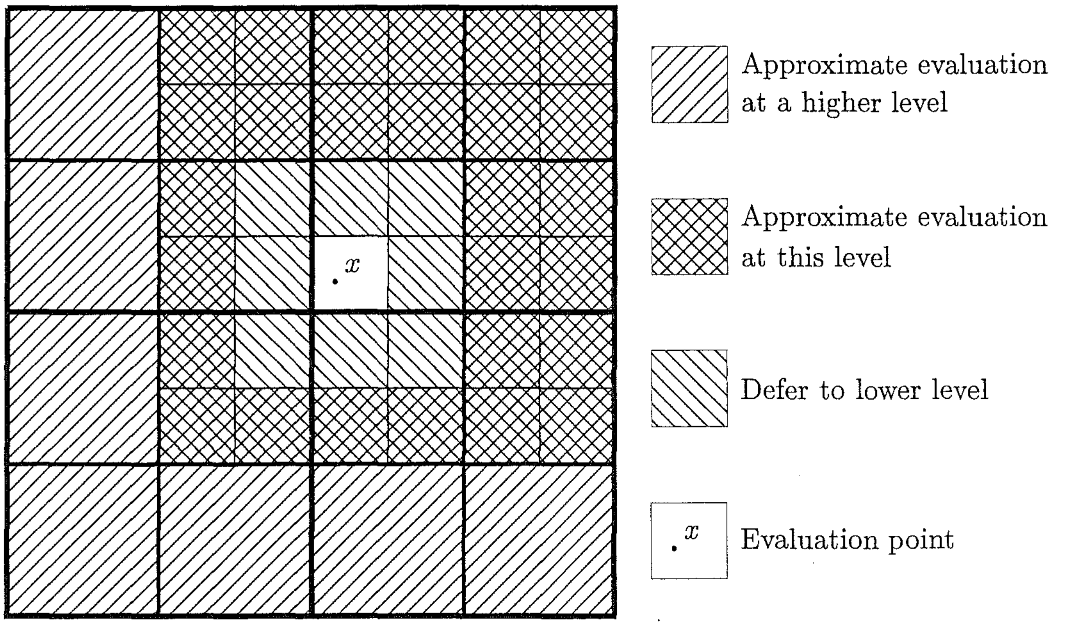


Figure 1.8: Evaluation of an RBF of two variables using the FMM based on a uniform quad-tree division of space. Each square is hatched based on how the part of the RBF due to sources in a given square is to be evaluated. The varying thickness of the lines that form the grid indicate the boundaries of the (square shaped) panels at various levels; the thicker the line the higher the level of the panel. In the case that there is no lower level, evaluation of that part of the RBF due to sources in the panels hatched ▧ must be done directly, *i.e.*, using the standard form of the RBF.

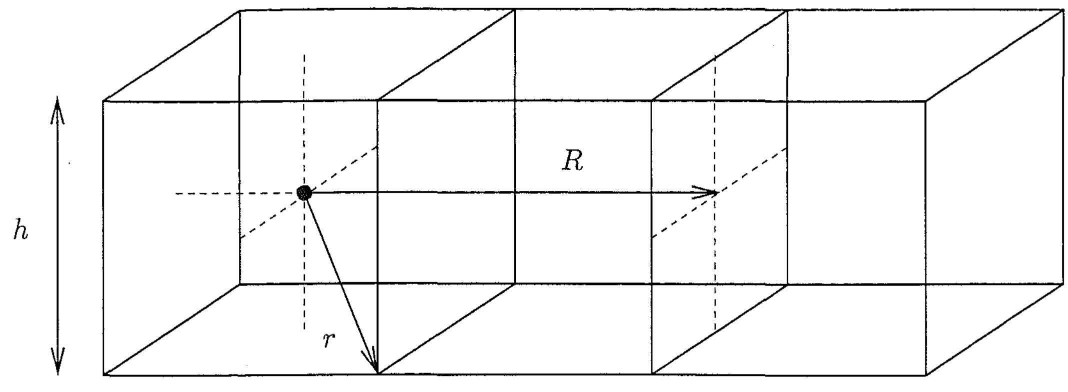


Figure 1.9: Ratio of distances of diagonal of cube to nearest interaction list member.

evaluator as described. Let the panels at a given level be cubes with sides of length h. Then $r = \sqrt{d(h/2)^2} = \sqrt{d}h/2$ and $R = 3h/2$ (see Figure 1.9). Thus $r/R = \sqrt{d}/3$. It now follows that if $d \geq 9$ then for a given level of precision it can not be guaranteed that there is a choice of p, and thus a corresponding truncated far field approximation r_T, such that $|s_T(x) - r_T(x)|$ is sufficiently small. In these higher dimensional cases a different strategy for dividing space must be used, or, at the very least, a different method for determining the suitability of approximate evaluation of a given source panel.

1.2.2 Improving setup cost for the hierarchical evaluator

The setup cost for this hierarchical evaluator is mainly in the computation of the coefficients

$$C_{T,\ell} = \sum_{i=1}^{N} \lambda_i P_\ell(x_i^2 + c^2, -2x_i, 1), \qquad \ell = 0, \ldots, p+1,$$

(or their equivalent). The cost of calculating these coefficients can be reduced by using recurrence relations and by translating child expansions to the parental panel.

Direct computation of these coefficients using Equation (3.4) would require $\mathcal{O}(p^2)$ flops to calculate P_ℓ, $\ell = 0, \ldots, p+1$. However, the polynomials P_ℓ may be defined in an equivalent recursive manner (see Lemma 3.5). Using this recurrence, the polynomials P_ℓ, $\ell = 0, \ldots, p+1$, may be calculated at a cost of $\mathcal{O}(p)$ flops. The inner and outer functions that play an equivalent role to the polynomials $P_\ell^{(k)}$ in the far field expansions for the polyharmonic splines in $\mathbb{R}^4$ also have a recursive definition (see Lemma 2.13 and Lemma 2.14).

Direct computation of the coefficients $C_{T,\ell}$ is only required at the finest level. At any other level in the tree, the coefficients for a particular panel T may be calculated indirectly and, in general, more efficiently from the coefficients of the children of T. This process is called *translation* of the far field expansion.

The name translation is used because of the way this indirect computation is done. Consider a panel T with two children, T_1 and T_2, say. Associated with each of these child panels is a truncated far field approximation (r_{T_1} and r_{T_2} respectively) to the appropriate parts of the RBF (s_{T_1} and s_{T_2}). Since $T_1 \cup T_2$ covers all of the source points in T,

$$s_T = s_{T_1} + s_{T_2}.$$

Furthermore, if $B_{T_1} \cup B_{T_2} \subseteq B_T$, then $r_{T_1}(x)+r_{T_2}(x)$ is an approximation to $s_T(x)$ that can be designed to be within any desired accuracy for any x outside of B_T. However, $r_T \neq r_{T_1} + r_{T_2}$ since r_T is a truncated far field expansion centred on the centre point of T while r_{T_1} and r_{T_2} are not; they are far field expansions but they are expanded about centre of T_1, T_2.

This leads to the idea of *translating* a far field expansion. Given a far field expansion g with centre u_1, find the far field expansion of g centred on u_2. Without loss of generality, it may be assumed that u_2 is in fact the origin and in this case the subscript on u_1 will be dropped. For the multiquadric, a far field expansion based around the polynomials $P_\ell^{(k)}$ may be translated by a convolutional approach described in Section 3.6. For the ordinary multiquadric of one variable if the technique described in that section is used, a far field expansion truncated to order p may be translated in $\mathcal{O}(p^2)$ flops. Similarly, a far field expansion for the polyharmonic RBF on $\mathbb{R}^4$ may be translated using Theorem 2.35.

The uniqueness results of Section 2.4 and Section 3.3 imply that two far field series, the one calculated directly from the source points and the one calculated indirectly from the expansions of the child panels, are actually identical. The immediate corollary of this is that no accuracy is lost by forming a far field series indirectly rather than directly.

1.2.3 The Fast Multipole Method

The Fast Multipole Method is an extension of the hierarchical evaluator described above; it reduces the cost of a single extra evaluation to $\mathcal{O}(1)$ flops. This is achieved by approximating all of the far field expansions for a given evaluation panel with a polynomial. This polynomial is based on truncated Taylor series approximations (centred on the centre of the evaluation panel) to each of the far field series. This polynomial is referred to as the *local expansion.* Like the far field series, the cost of evaluating a local expansion is dependent only on the accuracy required, not on the number of underlying source points or centres.

Calculating the Taylor series of a far field expansion, and thus the local expansion, is very similar to translating a far field series. For the multiquadric RBFs, the local series can be calculated via convolution in much the same way that the far field series was translated (see Section 3.7). Furthermore, in the 4D polyharmonic case exactly the same theorem, Theorem 2.35 is used to calculate the local series as is used to translate the far field series.

In much the same way that the far field series for the parent panels can be calculated from the far field series of the child panels, the local series of the parent panels may be inherited by the child panels. Once again, this is done by translating the centre of expansion from the centre of one panel to the centre of the other. Since the local series is just a polynomial, this is an exact recentering. Since the local series for the 4D polyharmonic RBFs are in terms of a special basis, the inner functions, Theorem 2.36 and related results are used to translate the local expansion and keep the nice basis. The local expansions for the multiquadric are in terms of the monomial basis and thus need no special algorithm.

With the FMM, for a particular evaluation panel T_x, the RBF is approximated by

$$s(x) \approx \sum_{\substack{T \,:\, x \in B_T \\ T \text{ is childless}}} s_T(x) + \text{polynomial}, \tag{1.39}$$

where the polynomial is the local expansion for T_x. For uniformly distributed data there will be $\mathcal{O}(1)$ terms in the sum in Equation 1.39 and thus the cost of a single evaluation by the FMM is $\mathcal{O}(1)$ flops, with an order constant that depends on the required accuracy. The cost of setup for the FMM is more than that for the hierarchical evaluator described above, but is still $\mathcal{O}(\log N)$ flops. Thus a matrix-vector product corresponding to evaluation at all centres can be computed in $\mathcal{O}(N \log N)$ flops using the FMM. For uniformly distributed data the only part of the FMM that is of $\mathcal{O}(N \log N)$ complexity is the sorting of centres into panels. The forming of all the series expansions and the evaluation has a total complexity of $\mathcal{O}(N)$. In contrast the evaluation part of any hierarchical method has complexity $\mathcal{O}(N \log N)$. Hence, in general, when N is large the FMM will be much faster than competing hierarchical methods.

1.2.4 Requirements of the Fast Multipole Method

There are two main aspects to the Fast Multipole Method: geometric and analytic. The geometric aspects involve how to decompose the domain into panels and conditions such as whether two panels are well separated. These implementation dependent geometric aspects are not the focus of this disertation and will not be considered any further. However, for a particular choice of basis function, there are a number of analytical results required that are independent of implementation. These required results are listed below.

- a *far field expansion*, with truncation error bound, for the basic function separating the influence of a *source* and *evaluation* point, *e.g.,* Theorem 2.24 for the potential in $\mathbb{R}^4$, Theorem 2.32 for a general polyharmonic spline in $\mathbb{R}^4$, and Theorem 3.9 for the generalised multiquadric in $\mathbb{R}^n$.
- an efficient means of calculating the coefficients in the far field expansion, *e.g.,* recursion relations such as those in Lemma 2.13, Lemma 2.14, and Lemma 3.5.
- a *uniqueness theorem* which shows that the truncated far field expansion is independent of the way in which it is calculated, *e.g.,* Lemma 2.33 and Lemma 2.34 for polyharmonic splines in $\mathbb{R}^4$, and Lemma 3.10 for the generalised multiquadric in $\mathbb{R}^n$.
- a means of *translating* a far field expansion to a different centre of expansion, *e.g.,* Theorem 2.35, Section 3.6.2.
- for the full FMM, a means for converting the far field expansion to a *local Taylor* or *local* series expansion, *e.g.,* Section 3.7.
- if the local Taylor series is in a special form, then a means of translating these series is also required, *e.g.,* Theorem 2.36.

1.3 "Theory and Application"?

The analytical results of Subsection 1.2.4 constitute the "theory" of fast evaluation of radial basis functions. Chapter 2 covers this theory for polyharmonic radial basis functions in four dimensions and Chapter 3 covers this theory for multiquadric radial basis functions in arbitrary dimensions. An application of the fast evaluation scheme for radial basis functions is the subject of Chapter 4. In this chapter RBFs are applied to the problem of fitting a surface to a cloud of points. This requires a subtle change in the apparent interpolation problem, but the interpolation problem that is finally solved is Problem 1.1. However, this is only one step in a two or three step process. The ability of an RBF to interpolate scattered data and the ability of the fast multipole method to efficiently evaluate large RBFs is what motivated the application of RBF to this particular surface fitting problem.

Chapter 2

Polyharmonic splines in $\mathbb{R}^4$

Polyharmonic splines in $\mathbb{R}^4$ are radial basis functions of the form

$$s = p + \sum_{k=1}^{N} d_k \phi(|\cdot - x_k|), \tag{2.1}$$

where the basic function ϕ is a member of the list

$$\phi(r) = \begin{cases} r^{-2}, \\ r^{2n} \ln(r), \quad n = 0, 1, \ldots, \end{cases} \tag{2.2}$$

and p is a low degree polynomial. A spline of this form is a solution to an iterated Laplace's equation, *i.e.*,

$$\Delta^{n+1} s = 0, \qquad \text{(a.e.)}.$$

These splines are the solutions in four dimensions to Problem 1.3. That is, interpolatory splines of this type minimize suitable energy semi-norms. For example, the functional

$$I(s) = \int_{\mathbb{R}^4} \sum_{|\alpha|=3} \binom{3}{\alpha} \Big((D^\alpha s)(x) \Big)^2 dx$$

is minimized over all suitably smooth functions, satisfying the interpolation conditions

$$s(x_k) = f(x_k), \qquad k = 1, \ldots, N, \tag{2.3}$$

if and only if s is the triharmonic spline

$$s = p_2 + \sum_{k=1}^{N} d_k |\cdot - x_k|^2 \ln |\cdot - x_k|,$$

where $p_2 \in \pi_2(\mathbb{R}^4)$, and the coefficients $\{d_k\}$, and those of the polynomial p_2, are determined by the interpolation conditions (2.3) along with the orthogonality conditions

$$\sum_{k=1}^{N} d_k q(x_k) = 0, \qquad \text{for all } q \in \pi_2(\mathbb{R}^4).$$

Thus polyharmonic splines in $\mathbb{R}^4$ can be expected to be highly successful approximators and interpolators, as experience has shown the polyharmonic splines in lower dimensions to be. However, meaningful data sets in $\mathbb{R}^4$ can be expected to have many points. Hence, the development of fast evaluation and fitting methods is almost a prerequisite to the use of polyharmonic splines in $\mathbb{R}^4$. Motivated by this we will develop the analytic underpinnings (see page 38) of a fast hierarchical and a fast multipole like method for polyharmonic splines in $\mathbb{R}^4$. There are many potential applications of these fast methods. One possible application to data mining is estimating the probability of some attribute, such as early death due to heart attack or the filing of a fraudulent tax return, by a regression spline depending on four predictor variables. An application to environmental engineering is modelling the concentration of some chemical, or pollutant, as a function of position and time. Turk & O'Brien [66] suggest using polyharmonic splines in $\mathbb{R}^4$ for shape transformation, or morphing, of implicitly defined surfaces in $3D$.

A significant technique in our development in this chapter is the use of a group action perspective. In particular, the use of arguments based on the action of the group of non-zero quaternions, realized as 2×2 complex matrices

$$\mathbb{H}_0^1 = \left\{ x = \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} : |z|^2 + |w|^2 > 0 \right\},$$

acting on $\mathbb{C}^2 = \mathbb{R}^4$. We develop almost all the (simple) details needed for these arguments without relying on other presentations of the possibly unfamiliar group representation theory. Use of this perspective allows us to give a relatively efficient development of the relevant spherical harmonics and their properties. See [55, 56] for related analyses of spherical harmonics and their approximation properties. Our work has also been influenced by the elegant and concise treatment by Epton & Dembart [25] of the analogous expansions for the three-dimensional fast multipole method.

This chapter is organised as follows. Section 2.1 concerns some of the properties of polyharmonic functions on $\mathbb{R}^4$—including realisations of $\mathbb{R}^4$ and representations of $\mathbb{H}_0^1$. It also introduces the inner and outer functions (spherical harmonics) that form the basis of our far field expansions. Section 2.2 develops a number of properties of these functions that can be applied to far field expansions. These include recurrence formulae, derivative formulae and symmetries. Section 2.3 contains the main results on the far field expansions themselves and the associated error bounds. Section 2.4 developes the uniqueness results that allow the far field expansions to be computed indirectly and economically via the translation theory of Section 2.5. Section 2.5 also contains the outer-to-inner and inner-to-inner translation formulae needed to approximate far field series by local Taylor series.

2.1 Polyharmonic functions on $\mathbb{R}^4$

We will represent a non-zero $x \in \mathbb{R}^4$ in three different ways:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \quad \text{or} \quad [z, w] \quad \text{or} \quad \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix},$$

where $z = x_1 + \mathrm{i}x_2$, $w = x_3 + \mathrm{i}x_4$ and $x_1, \ldots, x_4 \in \mathbb{R}$. The first realization is as an element of $\mathbb{R}^4$, the second is as an element of $\mathbb{C}^2$ and the last as an element of the punctured quaternion (Hamiltonian) space

$$\mathbb{H}_0^1 = \left\{ x = \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} : |z|^2 + |w|^2 > 0 \right\}. \tag{2.4}$$

Note that for elements of $\mathbb{H}_0^1$ the classical adjoint or adjugate and the Hermitian adjoint coincide and

$$x^{-1} = x^* / \det(x).$$

We are primarily interested in $\mathbb{R}^4$ with the usual inner product,

$$\langle x, x_< \rangle = x_1 x_{<1} + x_2 x_{<2} + x_3 x_{<3} + x_4 x_{<4} = |x||x_<| \cos\theta,$$

where $|\cdot|$ is the 2-norm for $\mathbb{R}^4$. In terms of the $\mathbb{C}^2$ realisation of $\mathbb{R}^4$ this becomes

$$\langle x, x_< \rangle = \Re(z\overline{z_<} + w\overline{w_<}) = \tfrac{1}{2}(z\overline{z_<} + w\overline{w_<} + z_<\overline{z} + w_<\overline{w})$$

and, in terms of the matrix realisation $\mathbb{H}_0^1$,

$$\langle x, x_< \rangle = \tfrac{1}{2}\operatorname{Tr}(x^* x_<) = \tfrac{1}{2}|x|^2 \operatorname{Tr}(x^{-1} x_<) = \tfrac{1}{2}\operatorname{Tr}(x_<{}^* x), \tag{2.5}$$

where Tr is the trace. Note that this inner product gives the norm

$$|x|^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 = z\overline{z} + w\overline{w} = \det(x).$$

We also require an inner product for functions on the unit ball $B = \{x \in \mathbb{R}^4 : |x| \leq 1\}$. For $f, g \in L^2(B)$ we define their inner product by

$$(f, g) = \int_{|\xi|\leq 1} f(\xi)\overline{g(\xi)}d\xi. \tag{2.6}$$

We will also use this pairing for other functions f and g on B with $f\overline{g} \in L^1(B)$. Furthermore, we will also require the subspaces of $C(\mathbb{R}^4)$ defined by

$$\mathcal{H}_m = \{p([z_1, z_2]) : p \text{ is a homogeneous polynomial of degree } m \text{ in } z_1, z_2\}, \qquad m \in \mathbb{N}_0,$$

where $\mathbb{N}_0$ is the non-negative integers.

Definition 2.1. A *polyharmonic* function f is a solution to the iterated Laplace's Equation

$$\Delta^k f = 0,$$

where k is a positive integer. To emphasize the order of polyharmonicity, f is also refered to as a k-harmonic function. For $k = 1, 2, 3$ the terms harmonic, biharmonic, and triharmonic, respectivily, are used.

2.1.1 Irreducible representations of $\mathbb{H}_0^1$–spherical harmonics

In this section we will develop irreducible representations of $\mathbb{H}_0^1$ in $\mathcal{H}_m$. Our purpose for doing so is that the coefficient functions of these representations will eventually be seen to form a very computationally convenient basis for the harmonic polynomials in $\mathbb{R}^4$. In fact when these coefficient functions are multiplied by $|x|^{2\ell}, \ell = 0, \ldots, k$ they yield a basis for all $(k+1)$-harmonic polynomials in $\mathbb{R}^4$. We note in particular the simple form of the addition formulae (Lemma 2.10), the recurrence relation (Lemma 2.13), and the dual basis (Lemma 2.22), to come.

Most of the relevant representation theory and mathematical physics literature is focused on rotations of $\mathcal{S}^2$ or $\mathcal{S}^3$ and therefore considers representations of

$$SU(2) = \left\{x \in \mathbb{H}_0^1 \,:\, \det(x) = |x|^2 = 1\right\} = \mathcal{S}^3.$$

However, in the context of far field expansions it is convenient to work instead with all of $\mathbb{H}_0^1$ to take into account both the scaling by $|x|$ and rotation by elements of $SU(2)$. This leads to some differences, most importantly the character functions now depend on the norm of x as well as the angle θ between x and the north pole $[1, 0]$. Other differences include the formula for the product of two character functions and the addition formulae.

Definition 2.2. Given a group G, a *representation* $T : G \to \mathrm{GL}(V)$ of G on V is an operator valued map that satisfies $T(g \cdot h) = T(g)T(h)$. A representation T of G on V is *irreducible* if the only subspaces of V that are invariant under $T(g)$ for all $g \in G$ are $\{0\}$ and V.

We define representations $T_m(x)$ of $\mathbb{H}^1_0$ in the spaces $\mathcal{H}_m$ given by

$$\begin{aligned} T_m(x)\ p([z_1, z_2]) &= p([z_1, z_2]x) \\ &= p([z_1 z - z_2\overline{w},\ z_1 w + z_2\overline{z}]). \end{aligned} \tag{2.7}$$

Note that since $\mathcal{H}_m$ is embedded in the space of functions on $\mathbb{H}^1$ (or even on $\mathbb{H}^1_0$), this representation is just the restriction of the right action defined by

$$(x \cdot f)([z_1, z_2]) = f([z_1, z_2]x), \quad x \in \mathbb{H}^1_0.$$

If we put a (Hilbert space) norm on these functions via (2.6), *i.e.*,

$$\|f\|^2 = (f, f) = \int_{|\xi|\le 1} |f(\xi)|^2 d\xi,$$

then the rotation invariance of Lebesgue measure implies that $\|x \cdot f\| = \|f\|$, whenever $|x| = 1$, *i.e.*, whenever $x \in \mathcal{S}^3$. Thus

$$\big(T_m(x)f, T_m(x)f\big) = (f, f)$$

for all functions $f \in L^2(B)$, and $T_m(x)$, as an operator, is unitary if $|x| = 1$. The reader is cautioned that the matrix realisation of $T_m(x)$ to come is not unitary, but can be scaled to be so (see Lemma 2.6).

Lemma 2.3. *The representations* (2.7) *are irreducible.*

Proof. Assume there is a subspace $V \subset \mathcal{H}_m$ that is invariant under $T_m(x)$ for all $x \in \mathbb{H}^1_0$. Then V is invariant under $T_m(x)$ if we restrict attention to $x \in SU(2) \subset \mathbb{H}^1_0$. Since the representations T_m of $SU(2)$ are irreducible [53, pp. 208–211], it follows that either $V = \{0\}$ or $V = \mathcal{H}_m$ and hence that the representations T_m are irreducible. ∎

The monomials

$$e^m_k([z_1, z_2]) = z_1^{m-k} z_2^k, \qquad 0 \le k \le m,$$

form a basis for $\mathcal{H}_m$. The operators $T_m(x)$ have a matrix realisation once this basis for $\mathcal{H}_m$ is chosen. The elements of these matrices will be denoted $t^m_{i,j}$ and are given by

$$\begin{aligned} t^m_{i,j}(x) &= t^m_{i,j}(z,w) \\ &= \text{coefficient of } e^m_i \text{ in } T_m(x)e^m_j \\ &= \text{coefficient of } z_1^{m-i}z_2^i \text{ in } (z_1 z - z_2\overline{w})^{m-j}(z_1 w + z_2\overline{z})^j. \end{aligned}$$

Equivalently,

$$\sum_{i=0}^{m} t^m_{i,j}(z,w) z_1^{m-i} z_2^i = (z_1 z - z_2\overline{w})^{m-j}(z_1 w + z_2\overline{z})^j. \tag{2.8}$$

For $m = 0$, $t^0_{0,0}(x) = 1$, while for $m = 1$, from (2.8)

$$\begin{aligned} t^1_{0,0}(z,w) &= z, & t^1_{0,1}(z,w) &= w, \\ t^1_{1,0}(z,w) &= -\overline{w}, & t^1_{1,1}(z,w) &= \overline{z}, \end{aligned}$$

or in matrix terms

$$[T_1(x)] = x. \tag{2.9}$$

An immediate consequence of this choice of basis is the following lemma.

Lemma 2.4. *Treated as matrices, $T_m(z,0)$ is a diagonal matrix and $T_m(0,w)$ is an anti-diagonal matrix. Specifically these matrices have entries*

$$t^m_{i,j}(z,0) = \begin{cases} 0, & i \neq j, \\ z^{m-i}\overline{z}^i, & i = j, \end{cases}$$

and

$$t^m_{i,j}(0,w) = \begin{cases} 0, & i \neq m-j, \\ w^{m-i}(-\overline{w})^i, & i = m-j. \end{cases}$$

Proof. Taking Equation (2.8) as the definition of the inner functions, $t^m_{i,j}(z,0)$ is given by

$$\sum_{i=0}^{m} t^m_{i,j}(z,0) z_1^{m-i} z_2^i = (z_1 z)^{m-j}(z_2\overline{z})^j.$$

Equating coefficients of $z_1^{m-i}z_2^i$ gives the first part of the lemma. Similary, $t^m_{i,j}(0,w)$ is given by

$$\sum_{i=0}^{m} t^m_{i,j}(0,w) z_1^{m-i} z_2^i = (-z_2\overline{w})^{m-j}(z_1 w)^j$$

and equating coefficients of $z_1^{m-i}z_2^i$ gives the second result. ∎

The basis elements e_k^m for $\mathcal{H}_m$ are orthogonal with respect to the inner product (2.6). In fact the exact norms for the basis elements e_k^m and thus the row and column scalings to get unitary matrices are easily computed.

Lemma 2.5. *The basis functions e_k^m are orthogonal with inner products given by*

$$(e_k^m, e_{k'}^{m'}) = \int_{|\xi|\leq 1} e_k^m(\xi)\overline{e_{k'}^{m'}(\xi)}d\xi = \delta_{m,m'}\delta_{k,k'}\pi^2\frac{k!(m-k)!}{(m+2)!}.$$

Proof. Introduce polar coordinates (r_1, θ_1) and (r_2, θ_2) in the z_1 and z_2 planes (where $\xi = [z_1, z_2]$).

$$\begin{aligned}
(e_k^m, e_{k'}^{m'}) &= \int_{|\xi|\leq 1} e_k^m(\xi)\overline{e_{k'}^{m'}(\xi)}d\xi \\
&= \int_0^1 \int_0^{\sqrt{1-r_1^2}} \int_0^{2\pi} \int_0^{2\pi} r_1^{m-k}e^{\mathrm{i}(m-k)\theta_1}r_2^k e^{\mathrm{i}k\theta_2} \\
&\qquad \times r_1^{m'-k'}e^{-\mathrm{i}(m'-k')\theta_1}r_2^{k'}e^{-\mathrm{i}k'\theta_2}\, d\theta_2 d\theta_1 r_2 dr_2 r_1 dr_1 \\
&= (2\pi)^2\delta_{k,k'}\delta_{m,m'}\int_0^1\int_0^{\sqrt{1-r_1^2}} r_1^{2(m-k)+1}r_2^{2k+1}dr_2dr_1 \\
&= (2\pi)^2\delta_{k,k'}\delta_{m,m'}\int_0^1 r_1^{2(m-k)+1}\frac{(1-r_1^2)^{k+1}}{2k+2}dr_1 \\
&= \pi^2\delta_{k,k'}\delta_{m,m'}\mathrm{B}(m-k+1,k+2)/(k+1) \\
&= \pi^2\delta_{k,k'}\delta_{m,m'}\frac{(m-k)!(k+1)!}{(m+2)!(k+1)},
\end{aligned}$$

where B is the Beta function $\mathrm{B}(n,m) = \Gamma(n)\Gamma(m)/\Gamma(n+m)$, see [1, § 6.2]. ∎

Since T_m restricted to $\mathcal{S}^3$ acts in a norm preserving way on $\mathcal{H}_m$, we easily obtain the following matrix representation for $T_m(x^{-1})$.

Lemma 2.6. *There exist row and column scalings that make the matrices $T_m(x)$ unitary for $|x| = 1$. Specifically*

(i) *The inverse of $T_m(x)$ is given by*

$$T_m(x^{-1}) = \left[t_{i,j}^m(x^{-1})\right] = \left[|x|^{-2m}\overline{t_{j,i}^m(x)}\binom{m}{i}\binom{m}{j}^{-1}\right],$$

or equivalently via

$$t_{i,j}^m(x^*) = t_{i,j}^m(\overline{z}, -w) = \overline{t_{j,i}^m(z,w)}\binom{m}{i}\binom{m}{j}^{-1}.$$

(ii) *For all $x \neq 0$, the inverses of the matrices*

$$U_m(x) = \left[t^m_{i,j}(x) \sqrt{\binom{m}{j}\binom{m}{i}^{-1}} \right]$$

are given by $U_m(x)^{-1} = |x|^{-2m} U_m(x)^*$ and thus are unitary when $|x| = 1$.

Proof. The definition of $t^m_{j,i}$ and the orthogonality of $\{e^m_k : k = 0, \ldots, m\}$ implies

$$(e^m_j, T_m(x) e^m_i) = (e^m_j, t^m_{j,i}(x) e^m_j) = \overline{t^m_{j,i}(x)} (e^m_j, e^m_j). \tag{2.10}$$

Since $T_m\big((x/|x|)^{-1}\big)$ preserves the inner product (2.6) and since T_m is homogeneous of degree m,

$$\begin{aligned}(e^m_j, T_m(x) e^m_i) &= \Big(T_m\big((x/|x|)^{-1}\big) e^m_j, T_m\big((x/|x|)^{-1}\big) T_m(x) e^m_i\Big) \\ &= \big(|x|^m T_m(x^{-1}) e^m_j, |x|^m e^m_i\big) = |x|^{2m} t^m_{i,j}(x^{-1}) (e^m_i, e^m_i). \end{aligned} \tag{2.11}$$

Equating (2.10) to (2.11) and solving, we obtain

$$t^m_{i,j}(x^{-1}) = |x|^{-2m} \overline{t^m_{j,i}(x)} \frac{(e^m_j, e^m_j)}{(e^m_i, e^m_i)}.$$

Using Lemma 2.5 to simplify $(e^m_j, e^m_j)/(e^m_i, e^m_i)$ gives part (i).

For part (ii) define the diagonal matrix

$$D_m = \begin{bmatrix} \sqrt{\binom{m}{0}} & & & \\ & \sqrt{\binom{m}{1}} & & \\ & & \ddots & \\ & & & \sqrt{\binom{m}{m}} \end{bmatrix}.$$

Since D_m is an invertable diagonal matrix , its inverse ${D_m}^{-1}$ is easily found and thus we may write

$$\begin{aligned} T_m(x)^{-1} &= |x|^{-2m} {D_m}^2 T_m(x)^* {D_m}^{-2}, \\ U_m(x) &= {D_m}^{-1} T_m(x) D_m. \end{aligned}$$

It now follows that

$$\begin{aligned} |x|^{-2m} U_m(x)^* U_m(x) &= |x|^{-2m} {D_m}^{-1} D_m U_m(x)^* U_m(x) \\ &= |x|^{-2m} {D_m}^{-1} {D_m}^2 T_m(x)^* {D_m}^{-2} T_m(x) D_m \\ &= {D_m}^{-1} T_m(x)^{-1} T_m(x) D_m \\ &= {D_m}^{-1} D_m = I. \end{aligned}$$

Thus $U_m(x)^{-1} = |x|^{2m} U_m(x)^*$ and furthermore $U_m(x)$ is unitary when $|x| = 1$. ∎

Since it is easy to use the definitions to show the first row of each $T_m(x)$ is given by

$$t_{0,j}^m(x) = t_{0,j}^m(z,w) = z^{m-j}w^j = e_j^m([z,w]),$$

part of Lemma 2.5 shows that $\{t_{0,j}^m\}$ is orthogonal with respect to the inner product (2.6). In fact much more general (bi-)orthogonality facts are true for $t_{i,j}^m(x)$ and $\overline{t_{j,i}^m(x^{-1})}$. These are related to the orthogonality properties of the irreducible unitary matrix representations of any compact group, such as $\mathcal{S}^3$, as in [35, (27.19)]. But we prefer to present them in a slightly more general form which is closely related to the coordinate free proofs in [2, Chapter 3], particularly Proposition 3.15, Schur's Lemma, 3.22, and its corollary, 3.23.

Lemma 2.7.

(i) *(Schur's Lemma) For any* $(m+1)\times(m+1)$ *matrix* A

$$\widetilde{A} = \int_{0<|x|\le 1} T_m(x^{-1})AT_m(x)dx = cI, \quad c = \frac{\mathrm{vol}\{|x|\le 1\}}{m+1}\operatorname{Tr}(A) = \frac{\pi^2/2}{m+1}\operatorname{Tr}(A).$$

(ii) *The set*

$$\left\{\frac{(m+1)}{(\pi^2/2)}\binom{m}{j}\binom{m}{i}^{-1}|\cdot|^{-2m}t_{i,j}^m(\cdot) = \frac{m+1}{\pi^2/2}\overline{t_{j,i}^m(\cdot^{-1})}, i,j=0,\dots,m\right\}$$

is biorthogonally dual to $\{t_{i,j}^m(\cdot), i,j=0,\dots,m\}$ *with respect to the pairing* (2.6). *That is,*

$$\int_{0<|x|\le 1}\frac{m+1}{\pi^2/2}t_{i',j'}^m(x)t_{j,i}^m(x^{-1})dx = \delta_{i,i'}\delta_{j,j'}.$$

(iii) *The first two parts are also true when* $\{0<|x|\le 1\}$ *is replaced by* $\mathcal{S}^3$ *and "vol" is replaced by "surface area" (so* $\pi^2/2$ *is replaced by* $2\pi^2$*).*

Proof. For (i), let $y\in\mathcal{S}^3$ be arbitrary. Then

$$\widetilde{A}T_m(y) = \int_{0<|x|\le 1} T_m(x^{-1})AT_m(xy)dx = \int_{0<|x|\le 1} T_m(yx^{-1})AT_m(x)dx = T_m(y)\widetilde{A}, \quad (2.12)$$

since $x\mapsto xy^{-1}$ leaves Lebesgue measure invariant. Let c be any eigenvalue for $\widetilde{A}$ with v an associated eigenvector. From (2.12) $T_m(y)v$ is also an eigenvector for the same eigenvalue c. By the irreducibility of T_m, span $\{T_m(y)v : y\in\mathcal{S}^3\} = \mathcal{H}_m$. Thus $\widetilde{A}v = cv$ for all vectors v and it follows that $\widetilde{A} = cI$.

To get the formula for c, take the trace of all terms. Then move the linear functional Tr inside the integral and use $\operatorname{Tr}(T_m(x^{-1})AT_m(x)) = \operatorname{Tr}(A)$ to obtain

$$\operatorname{Tr}(cI) = (m+1)c = \int_{0<|x|\le 1}\operatorname{Tr}(A)dx = \mathrm{vol}\{0<|x|\le 1\}\operatorname{Tr}(A).$$

For (ii) substitute $A = E_{i,i'}$ into (i), where $E_{i,i'}$ is the matrix with $\delta_{i,j}\delta_{j',i'}$ in the (j,j') position. Then

$$T_m(x^{-1})E_{i,i'}T_m(x) = [t^m_{j,i}(x^{-1})t^m_{i',j'}(x)].$$

Since $\operatorname{Tr} E_{i,i'} = \delta_{i,i'}$, (i) yields

$$\left[\int_{0<|x|\leq 1} t^m_{j,i}(x^{-1})t^m_{i',j'}(x)\,dx\right] = \frac{\pi^2/2}{m+1}\delta_{i,i'}[\delta_{j,j'}].$$

For (iii) repeat the proofs with the ball replaced by the sphere. Or, more simply, just note that if $d\Omega$ denotes the standard "surface" measure on $\mathcal{S}^3$, then integration in spherical coordinates is with respect to $|x|^3 d\Omega(x/|x|)d|x|$. Then, due to the homogeneity of T_m, (i) becomes

$$\begin{aligned}\int_{\mathcal{S}^3}\int_{|x|=0}^{|x|=1} & T_m\big((x/|x|)^{-1}\big)AT_m(x/|x|)|x|^3 d|x|d\Omega(x/|x|)\\ &= \frac{1}{4}\int_{\mathcal{S}^3} T_m\big((x/|x|)^{-1}\big)AT_m(x/|x|)d\Omega(x/|x|)\\ &= \frac{\pi^2/2}{m+1}\operatorname{Tr}(A).\end{aligned}$$

Clearing the denominator of 4 leads to the desired formula for (i) on $\mathcal{S}^3$. Now (ii) on $\mathcal{S}^3$ follows by exactly the same reasoning. Since $|x|^{-2m} = 1$ on $\mathcal{S}^3$, the orthogonality of $t^m_{i,j}$ on $\mathcal{S}^3$ follows, as does their independence since none of these functions are zero (or have norm zero on $\mathcal{S}^3$). ∎

Lemma 2.7(iii) implies that the coefficient functions are linearly independent in both $L^2(\mathcal{S}^3)$ and $C(\mathbb{R}^4)$. Indeed, they form a basis for homogeneous harmonic polynomials on $\mathbb{R}^4$ and for the spherical harmonics of degree m on $\mathcal{S}^3$ (see Equation (2.33)). Hence, for any $p \in \mathbb{N}$, $\{t^m_{i,j} : 0 \leq i,j \leq m, 0 \leq m \leq p\}$ is linearly independent both on $\mathcal{S}^3$ and on $\mathbb{R}^4$.

Given the north pole $[1,0]$ and some general vector $x = [z,w]$, we can find a rotation that leaves the north pole fixed and rotates x to a vector in the direction $[e^{i\theta},0]$ where $\cos\theta = \Re(z)/|x|$. Note that θ is just the angle between x and the north pole. We could equivalently rotate to a vector in the direction $[e^{-i\theta},0]$. Hence any function independent of rotation about the north pole must be a function of $|x|$ and θ and furthermore must be even in θ, *i.e.*, is a function of $\cos\theta$. It is known that all rotations leaving the north pole fixed can be achieved by conjugation, $x \mapsto vxv^{-1}$, with elements of $SU(2)$. See [41, pp. 277–279] or [3, pp. 214–217] for a geometric proof. The same result may be obtained algebraically by considering the diagonalizability of x (see, *e.g.*, [53, pp. 209–210]). Therefore there is a $v \in SU(2)$ such that

$$x = v\gamma v^{-1}$$

where

$$\gamma = |x| \operatorname{diag}(e^{\mathrm{i}\theta}, e^{-\mathrm{i}\theta}).$$

Note that conjugation with $-v$ achieves the same rotation as conjugation with v, but this is the only nonuniqueness in identifying conjugations with rotations of the equatorial $\mathcal{S}^2$.

These conjugation facts lead to explicit formulae for the traces of these representations. Specifically,

$$\begin{aligned}\operatorname{Tr}\left(T_m(x)\right) = \operatorname{Tr}\left(T_m(v\gamma v^{-1})\right) &= \operatorname{Tr}\left(T_m(v)T_m(\gamma)T_m(v^{-1})\right) \\ &= \operatorname{Tr}\left(T_m(v)T_m(\gamma)T_m(v)^{-1}\right) = \operatorname{Tr}\left(T_m(\gamma)\right)\end{aligned}$$

Using Lemma 2.4,

$$t^m_{i,j}(\gamma) = t^m_{i,j}(|x|e^{\mathrm{i}\theta}) = \begin{cases} |x|^m e^{\mathrm{i}(m-2j)\theta} & i = j, \\ 0 & i \neq j. \end{cases}$$

Hence

$$\begin{aligned}\operatorname{Tr}\left(T_m(x)\right) = \operatorname{Tr}\left(T_m(\gamma)\right) &= \sum_{j=0}^{m} t^m_{j,j}(\gamma) \\ &= \sum_{j=0}^{m} |\gamma|^m e^{\mathrm{i}(m-2j)\theta} \\ &= |\gamma|^m \frac{e^{\mathrm{i}(m+1)\theta} - e^{-\mathrm{i}(m+1)\theta}}{e^{\mathrm{i}\theta} - e^{-\mathrm{i}\theta}}, \qquad \text{if } \theta \neq 0, \\ &= |x|^m \frac{\sin(m+1)\theta}{\sin\theta},\end{aligned}$$

and interpreting $\sin(m+1)\theta/\sin\theta$ in the conventional fashion (as $m+1$) at $\theta = 0$, the expression is also valid there. As usual the character of the representation is defined to be the function $\chi_m : \mathbb{H}^1_0 \to \mathbb{R}$ given by the trace

$$\chi_m(x) := \operatorname{Tr}\left(T_m(x)\right) = |x|^m \frac{\sin(m+1)\theta}{\sin\theta}. \tag{2.13}$$

In particular $\chi_0(x) = 1$ and $\chi_1(x) = \operatorname{Tr}(x) = 2|x|\cos\theta$. These χ_m are homogeneous polynomials of degree m in z, w, $\overline{w}$ and $\overline{z}$ since the entries in $T_m(x)$ are. We extend the definition of χ_m to $x = 0$ by continuity and define $\chi_{-1} = 0$. Note that these χ_m are multiples of the Chebyshev polynomials of the second kind as functions of $t = \cos\theta$.

Lemma 2.8. *For $x \in \mathbb{H}^1$ and $m \in \mathbb{N}_0$,*

$$\chi_1(x)\chi_m(x) = \chi_{m+1}(x) + |x|^2\chi_{m-1}(x). \tag{2.14}$$

Proof. The result is trivially true when $m = 0$ or $x = 0$. For $m > 0$ and $x \neq 0$

$$\begin{aligned}
\chi_1(x)\chi_m(x) &= \frac{|x|^{m+1}}{\sin^2\theta}\Big\{\sin(2\theta)\sin\big((m+1)\theta\big)\Big\} \\
&= \frac{|x|^{m+1}}{\sin^2\theta}\Big\{2\sin(\theta)\cos(\theta)\sin\big((m+1)\theta\big)\Big\} \\
&= \frac{|x|^{m+1}}{\sin\theta}\Big\{\sin\big((m+2)\theta\big) + \sin(m\theta)\Big\} \\
&= \chi_{m+1}(x) + |x|^2\chi_{m-1}(x).
\end{aligned}$$

∎

2.1.2 Inner and outer functions

We will refer to the functions $t^m_{i,j}$ of the previous subsection as the *inner* functions as they will be shown to be homogeneous of non-negative degree and harmonic in $\mathbb{R}^4$. The purpose of this subsection is to introduce the *outer* functions $o^m_{i,j}$ which will be shown to be homogeneous of negative degree and harmonic in $\mathbb{R}^4\backslash\{0\}$. The subsection also contains the addition formula connecting the inner and outer functions with the character functions.

The representations T_m may be used to construct anti-representations, O_m of $\mathbb{H}^1_0$, defined by

$$O_m(x) := |x|^{-2}T_m(x^{-1}) \tag{2.15}$$

or, in terms of the coefficient functions,

$$\begin{aligned}
o^m_{i,j}(z,w) = |x|^{-2}t^m_{i,j}(x^{-1}) &= (z\overline{z} + w\overline{w})^{-1}t^m_{i,j}(\overline{z}/|x|^2, -w/|x|^2) \\
&= |x|^{-(2m+2)}t^m_{i,j}(\overline{z}, -w) \\
&= |x|^{-2(m+1)}\overline{t^m_{j,i}(z,w)}\binom{m}{i}\binom{m}{j}^{-1}.
\end{aligned} \tag{2.16}$$

Since $t^m_{j,i}$ is homogeneous of degree m, the coefficient functions for O_m are homogeneous of degree $-(m+2)$. Together, (2.8) and (2.16) give the equivalent definition of the outer functions

$$|x|^{2(m+1)}\sum_{i=0}^{m} o^m_{i,j}(z,w)z_1^{m-i}z_2^i = (z_1\overline{z} + z_2\overline{w})^{m-j}\big(z_1(-w) + z_2 z\big)^j. \tag{2.17}$$

See appendix A for tables of low degree inner and outer functions. The operators O_m are anti-representations as

$$O_m(x\cdot y) = O_m(y)\cdot O_m(x).$$

Remark 2.9. When we prove harmonicity of the inner functions, or of the outer functions, the harmonicity of the other set will follow. Indeed definition (2.16) corresponds to an inversion of the functions $t^m_{i,j}$ in the unit sphere followed by scaling by $|x|^{-2}$, along with reflection in the $\Re(z)$ axis. This reflection $(z, w) \to (\overline{z}, -w)$ corresponds to quaternionic conjugation. Both this scaled inversion, sometimes called the Kelvin transformation, and the reflection preserve harmonicity.

Associated with the spherical harmonics in $\mathcal{S}^{d-1}$ for any integer $d \geq 2$ is an addition formula [51, pp. 3–10]. These addition formulae express the character function (sometimes called the Gengenbauer polynomial or the Legendre function) at the inner product of two points u and v on $\mathcal{S}^{d-1}$ as a sum of products, in each of which the influence of u and v is separated. Perhaps the best known example of this phenomena is the addition formula for the ordinary Legendre polynomial, $P_n(\cos\gamma)$, which is exploited in the multipole expansion of the 3D potential, (see, *e.g.*, [9, 13, 25]). With our definition of the inner and outer functions, the addition formula for $\mathbb{R}^4$ takes the forms

Lemma 2.10 (Addition formulae for χ_m).

(i) *If* $x, x_< \in \mathbb{H}^1$, $x \neq 0$ *then*

$$\chi_m(x^{-1}x_<) = |x|^2 \operatorname{Tr}\left(O_m(x)T_m(x_<)\right) = |x|^2 \sum_{i,j=0}^{m} t^m_{j,i}(x_<)\, o^m_{i,j}(x). \tag{2.18}$$

(ii) *If* $x, x_< \in \mathbb{H}^1$ *then*

$$\chi_m(x^* x_<) = \operatorname{Tr}\left(T_m(x_<{}^*)T_m(x)\right) = \sum_{i,j=0}^{m} t^m_{j,i}(x_<{}^*)\, t^m_{i,j}(x). \tag{2.19}$$

Proof. From the definition of the character functions χ_m,

$$\begin{aligned}
\chi_m(x^{-1}x_<) = \operatorname{Tr}\left(T_m(x^{-1}x_<)\right) &= \operatorname{Tr}\left(T_m(x^{-1})T_m(x_<)\right) \\
&= \sum_{i,j=0}^{m} t^m_{i,j}(x^{-1}) t^m_{j,i}(x_<) \\
&= |x|^2 \sum_{i,j=0}^{m} |x|^{-2} t^m_{i,j}(x^{-1}) t^m_{j,i}(x_<) \\
&= |x|^2 \sum_{i,j=0}^{m} o^m_{i,j}(x) t^m_{j,i}(x_<),
\end{aligned}$$

which proves part (i).

Since the mapping $x \mapsto x^*$ corresponds to reflection across the axis of the north pole, the angle between x and the north pole is equal to the angle between x^* and the north pole. Thus, Equation (2.13) implies $\chi_m(x^*) = \chi_m(x)$ and therefore

$$\chi_m(x^* x_<) = \chi_m(x_<{}^* x) = \mathrm{Tr}\left(T_m(x_<{}^*) T_m(x)\right) = \sum_{i,j=0}^{m} t^m_{j,i}(x_<{}^*)\, t^m_{i,j}(x)$$

∎

The addition formula (2.18) essentially displays the fact that

$$\frac{m+1}{\pi^2/2}\chi_m(x^{-1}x_<)$$

is a reproducing kernel for $\mathrm{span}\{t^m_{i,j}, i,j = 0,\ldots,m\}$, the space of homogeneous harmonic polynomials of degree m. In fact the biorthogonality in Lemma 2.7(ii) immediately shows that for any

$$f_m = \sum_{i,j=0}^{m} a_{i,j} t^m_{i,j}$$

in this span,

$$\int_{0<|x|\leq 1} f_m(x) \frac{m+1}{\pi^2/2}\chi_m(x^{-1}x_<)dx = f_m(x_<).$$

2.2 Properties of the inner and outer functions

Some fundamental properties of the inner and outer functions will be developed in this section. These properties are needed for the development of a fast multipole like method, but are also of interest in their own right. Properties which are considered below include symmetries, recurrence relations, derivative formulae and harmonicity.

2.2.1 Symmetries

In this subsection we will develop some symmetry properties of the inner and outer functions. We have already seen an example of a symmetry relation in Lemma 2.6. One application of these symmetry properties is an approximate halving of the costs of forming and evaluating the truncated far field expansions to be developed in Section 2.3.

We will use the symbols i, j and k for the fundamental quaternionic units and $\mathbf{i}$ for the imaginary number $\sqrt{-1}$. It will be convenient to use the 2×2 matrix realisation of the quaternions. In this realisation the quaternion 1 is the 2×2 identity matrix and

$$\mathsf{i} = \begin{bmatrix} \mathbf{i} & 0 \\ 0 & -\mathbf{i} \end{bmatrix}, \qquad \mathsf{j} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad \mathsf{k} = \begin{bmatrix} 0 & \mathbf{i} \\ -\mathbf{i} & 0 \end{bmatrix}.$$

The basic relations between the quaternionic units are

$$\mathsf{i}^2 = \mathsf{j}^2 = \mathsf{k}^2 = -1,$$

$$\mathsf{ij} = \mathsf{k}, \quad \mathsf{jk} = \mathsf{i}, \quad \mathsf{ki} = \mathsf{j}.$$

Further properties of the quaternions may be found in, *e.g.*, [3, 41].

Now consider conjugation of an element in $\mathbb{H}^1$ by i.

$$\begin{bmatrix} \mathbf{i} & 0 \\ 0 & -\mathbf{i} \end{bmatrix} \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \begin{bmatrix} -\mathbf{i} & 0 \\ 0 & \mathbf{i} \end{bmatrix} = \begin{bmatrix} z & -w \\ \overline{w} & \overline{z} \end{bmatrix}.$$

Applying T_m

$$T_m(\mathbf{i}, 0) T_m(z, w) T_m(-\mathbf{i}, 0) = T_m(z, -w).$$

By Lemma 2.4 $T_m(\mathbf{i}, 0)$ and $T_m(-\mathbf{i}, 0)$ are diagonal with

$$t^m_{i,i}(\mathbf{i}, 0) = (-1)^i \mathbf{i}^m, \qquad t^m_{j,j}(-\mathbf{i}, 0) = (-1)^{m-j} \mathbf{i}^m.$$

We conclude that

$$(-1)^i \mathbf{i}^m t^m_{i,j}(z, w)(-1)^{m-j} \mathbf{i}^m = t^m_{i,j}(z, -w)$$

for $0 \leq i, j \leq m$. That is

$$t^m_{i,j}(z, w) = (-1)^{i-j} t^m_{i,j}(z, -w) \tag{2.20}$$

for $0 \leq i, j \leq m$.

Similar results can be obtain in the same way by conjugation with j and k. These results are summarised in the following lemma.

Lemma 2.11. *The inner functions $t^m_{i,j}$ defined by (2.8) satisfy*

$$t^m_{i,j}(z, w) = (-1)^{i-j} t^m_{i,j}(z, -w), \tag{2.21a}$$

$$t^m_{i,j}(z, w) = (-1)^{i+j} t^m_{m-i,m-j}(\overline{z}, \overline{w}), \tag{2.21b}$$

$$t^m_{i,j}(z, w) = t^m_{m-i,m-j}(\overline{z}, -\overline{w}), \tag{2.21c}$$

for all $x = [z, w] \in \mathbb{C}^2$.

Because each $t^m_{i,j}(z, w)$ is a polynomial with real coefficients in z, w, $\overline{z}$ and $\overline{w}$,

$$\overline{T_m(z, w)} = T_m(\overline{z}, \overline{w}).$$

Hence (2.21b) may be written

$$t^m_{i,j}(z, w) = (-1)^{i+j} \overline{t^m_{m-i,m-j}(z, w)}. \tag{2.22}$$

Using the expression (2.16) for the outer functions in terms of the inner functions each symmetry of T_m implies a symmetry of O_m. In particular, symmetry (2.22) implies

$$\begin{aligned} o^m_{i,j}(z,w) &= |x|^{-(2m+2)} t^m_{i,j}(\overline{z}, -w) \\ &= |x|^{-(2m+2)} (-1)^{i+j} t^m_{m-i,m-j}(z, -\overline{w}) \\ &= (-1)^{i+j} \overline{o^m_{m-i,m-j}(z,w)}. \end{aligned} \tag{2.23}$$

From symmetries (2.22) and (2.23)

$$o^m_{i,j}(x) t^m_{j,i}(x_<) = \overline{o^m_{m-i,m-j}(x) t^m_{m-j,m-i}(x_<)},$$

for $0 \le i,j \le m$. Hence the terms in the addition formula expression (2.18) for $\chi_m(x^{-1}x_<)$ are conjugate symmetric with respect to reflection in the middle index $(\frac{m}{2}, \frac{m}{2})$. Thus the addition formula can be rewritten to involve a real part and approximately half the number of terms. For example, one such recasting is

$$\begin{aligned} \chi_m(x^{-1}x_<) &= |x|^2 \sum_{i,j=0}^{m} o^m_{i,j}(x) t^m_{j,i}(x_<) \\ &= |x|^2 \Re \left(\sum_{k=0}^{\left\lfloor \frac{1}{2}\left((m+1)^2-1\right) \right\rfloor} (2 - \delta_{m/2,i(k)}\, \delta_{m/2,j(k)}) o^m_{i(k),j(k)}(x) t^m_{j(k),i(k)}(x_<) \right), \end{aligned} \tag{2.24}$$

where $i(k) = k \mod (m+1)$ and $j(k) = \lfloor k/(m+1) \rfloor$. This observation can and should be used to halve the storage requirements and flop counts of fast evaluators built upon the analysis of this chapter. For example, the recast expression (2.24) can be substituted almost directly into the truncated far field expansion, g_p, for a sum of shifts of the potential of Theorem 2.24. This will then approximately halve the marginal operation count for evaluating $g_p(x)$ at a single extra x.

A further symmetry will be useful to recast products of powers of $|x|$ with outer functions as inner functions.

Lemma 2.12. *For $m \in \mathbb{N}$ and $0 \le i,j \le m$*

$$\binom{m}{j} |x|^{2m+2} o^m_{i,j}(x) = (-1)^{i-j} \binom{m}{i} t^m_{m-j,m-i}(x).$$

Proof. Substitution of (2.22) into (2.16) gives the result. ∎

2.2.2 Recurrence relationships

In this subsection we develop recurrence relations which provide efficient methods for calculating the inner and outer functions. For a given x they allow calculation of $\{t_{i,j}^m,\ 0 \le i,j \le m,\ 0 \le m \le p\}$ or $\{o_{i,j}^m,\ 0 \le i,j \le m,\ 0 \le m \le p\}$ in $\mathcal{O}(p^3)$ operations, which is the same magnitude as the number of terms to be calculated.

It is convenient to extend the definitions of $t_{i,j}^m$ and $o_{i,j}^m$ by defining $t_{i,j}^m = 0 = o_{i,j}^m$ for i or $j \notin [0,m]$.

Lemma 2.13. *The inner functions $t_{i,j}^m$ defined by (2.8), satisfy the overlapping recurrence relations:*

$$t_{i,j}^{m+1}(z,w) = \begin{cases} z\, t_{i,j}^m(z,w) - \overline{w}\, t_{i-1,j}^m(z,w), & 0 \le j \le m, \\ \\ w\, t_{i,j-1}^m(z,w) + \overline{z}\, t_{i-1,j-1}^m(z,w), & 1 \le j \le m+1, \end{cases} \tag{2.25}$$

for $0 \le i \le m+1$, and also satisfy the backward recurrence relations:

$$\begin{aligned} |x|^2 t_{i,j}^m(z,w) &= \overline{z}\, t_{i,j}^{m+1}(z,w) + \overline{w}\, t_{i,j+1}^{m+1}(z,w) \\ &= -w\, t_{i+1,j}^{m+1}(z,w) + z\, t_{i+1,j+1}^{m+1}(z,w) \end{aligned} \tag{2.26}$$

for $0 \le i,j \le m$.

Proof. From (2.8), for $1 \le j \le m+1$, we obtain

$$\begin{aligned} \sum_{i=0}^{m+1} t_{i,j}^{m+1}(z,w)\; z_1^{m+1-i} z_2^i &= (zz_1 - \overline{w}z_2)^{m+1-j}(wz_1 + \overline{z}z_2)^j \\ &= \left\{(zz_1 - \overline{w}z_2)^{m-(j-1)}(wz_1 + \overline{z}z_2)^{j-1}\right\}(wz_1 + \overline{z}z_2) \\ &= \left\{\sum_{i=0}^{m} t_{i,j-1}^m(z,w)\; z_1^{m-i} z_2^i\right\}(wz_1 + \overline{z}z_2) \\ &= \sum_{i=0}^{m} t_{i,j-1}^m(z,w)\left\{wz_1^{m+1-i}z_2^i + \overline{z}z_1^{m-i}z_2^{i+1}\right\} \\ &= \sum_{i=0}^{m} t_{i,j-1}^m(z,w)wz_1^{m+1-i}z_2^i + \sum_{i=0}^{m} t_{i,j-1}^m(z,w)\overline{z}z_1^{(m+1)-(i+1)}z_2^{i+1} \\ &= \sum_{i=0}^{m} w\, t_{i,j-1}^m(z,w)z_1^{m+1-i}z_2^i + \sum_{i=1}^{m+1} \overline{z}\, t_{i-1,j-1}^m(z,w)z_1^{m+1-i}z_2^i \\ &= \sum_{i=0}^{m+1}\left\{w\, t_{i,j-1}^m(z,w) + \overline{z}\, t_{i-1,j-1}^m(z,w)\right\} z_1^{m+1-i}z_2^i. \end{aligned}$$

Equating coefficients of $z_1^{m+1-i}z_2^i$ gives the second part of (2.25). The first part may be obtained in a similar manner, or more directly by application of the symmetry relation (2.21b) to both sides of the part just proved.

The backward recursion may be obtained directly from the forward recursion. Multiplying the first right-hand side of (2.25) by $\overline{z}$ and the second right-hand side (with j replaced by $j+1$) by $\overline{w}$ and summing gives the first right-hand side of (2.26). The second right-hand side is obtained similarly. ∎

These formulae are analogous to the addition formulae which express $\cos(m \pm 1)\theta$ and $\sin(m \pm 1)\theta$ in terms of $\cos\theta$, $\sin\theta$, $\sin m\theta$ and $\cos m\theta$.

Lemma 2.14. *The outer functions $o_{i,j}^m$ defined by (2.16), satisfy the overlapping recurrence relations:*

$$o_{i,j}^{m+1}(z,w) = \begin{cases} \dfrac{\overline{z}}{|x|^2} o_{i,j}^m(z,w) + \dfrac{\overline{w}}{|x|^2} o_{i-1,j}^m(z,w), & 0 \le j \le m, \\[2ex] \dfrac{-w}{|x|^2} o_{i,j-1}^m(z,w) + \dfrac{z}{|x|^2} o_{i-1,j-1}^m(z,w), & 1 \le j \le m+1, \end{cases} \tag{2.27}$$

for $0 \le i \le m+1$, and also satisfy the backward recurrence relations:

$$\begin{aligned} o_{i,j}^m(z,w) &= z\, o_{i,j}^{m+1}(z,w) - \overline{w}\, o_{i,j+1}^{m+1}(z,w) \\ &= w\, o_{i+1,j}^{m+1}(z,w) + \overline{z}\, o_{i+1,j+1}^{m+1}(z,w) \end{aligned} \tag{2.28}$$

for $0 \le i,j \le m$.

Proof. This follows easily by substituting the expression (2.16) for $o_{i,j}^m$ into the relations (2.25) and (2.26) for the inner functions $t_{i,j}^m$. ∎

2.2.3 Derivatives and harmonicity

In this subsection we develop derivative formulae for the inner and outer functions. Applications include proofs of harmonicity, a dual basis for the inner functions, and an expression for the outer functions as an appropriate derivatives of $1/|x|^2$. Analogous formulae in the three dimensional case are given in Epton & Dembart [25].

We start by recalling the definitions for the complex derivative operators:

$$\begin{aligned} \frac{\partial}{\partial z} &= \frac{1}{2}\left(\frac{\partial}{\partial x_1} - \mathrm{i}\frac{\partial}{\partial x_2}\right), & \frac{\partial}{\partial \overline{z}} &= \frac{1}{2}\left(\frac{\partial}{\partial x_1} + \mathrm{i}\frac{\partial}{\partial x_2}\right), \\ \frac{\partial}{\partial w} &= \frac{1}{2}\left(\frac{\partial}{\partial x_3} - \mathrm{i}\frac{\partial}{\partial x_4}\right), & \frac{\partial}{\partial \overline{w}} &= \frac{1}{2}\left(\frac{\partial}{\partial x_3} + \mathrm{i}\frac{\partial}{\partial x_4}\right). \end{aligned} \tag{2.29}$$

By considering z, $\overline{z}$, w and $\overline{w}$ as functions of x_1, x_2, x_3 and x_4, these operators may be applied to functions defined on $\mathbb{C}^2$ in the natural way. Indeed, immediate consequences of these definitions are the relations

$$\frac{\partial}{\partial z}\overline{z} = 0, \qquad \text{and} \qquad \frac{\partial}{\partial z} z = 1,$$

their conjugates and the similar relations with w, $\overline{w}$. More generally, if $f(x) = h(z, \overline{z}, w, \overline{w})$ is complex analytic in z, *i.e.*, independent of $\overline{z}$ in the sense that $\partial h/\partial \overline{z} = 0$, then $\partial f/\partial z$ is given by all the usual rules for differentiation. Furthermore, armed with these operators the ordinary Laplacian may be expressed as

$$\Delta = 4\left(\frac{\partial^2}{\partial z \partial \overline{z}} + \frac{\partial^2}{\partial w \partial \overline{w}}\right). \tag{2.30}$$

Lemma 2.15. *Any of the first partial derivatives map the inner functions of degree m to (multiples of) inner functions of degree $m-1$. Specifically, for $0 \le i, j \le m+1$,*

$$\frac{\partial}{\partial z} t_{i,j}^{m+1}(z,w) = (m+1-j) t_{i,j}^{m}(z,w), \tag{2.31a}$$
$$\frac{\partial}{\partial \overline{w}} t_{i,j}^{m+1}(z,w) = -(m+1-j) t_{i-1,j}^{m}(z,w), \tag{2.31b}$$
$$\frac{\partial}{\partial w} t_{i,j}^{m+1}(z,w) = j t_{i,j-1}^{m}(z,w), \tag{2.31c}$$
$$\frac{\partial}{\partial \overline{z}} t_{i,j}^{m+1}(z,w) = j t_{i-1,j-1}^{m}(z,w), \tag{2.31d}$$

Proof. Differentiating (2.8) with respect to z gives

$$\begin{aligned}
\sum_{i=0}^{m+1} \frac{\partial}{\partial z} t_{i,j}^{m+1}(z,w) z_1^{m+1-i} z_2^i &= (m+1-j) z_1 (z_1 z - z_2 \overline{w})^{m-j} (z_1 w + z_2 \overline{z})^j \\
&= (m+1-j) z_1 \sum_{i=0}^{m} t_{i,j}^{m}(z,w) z_1^{m-i} z_2^i \\
&= (m+1-j) \sum_{i=0}^{m} t_{i,j}^{m}(z,w) z_1^{m+1-i} z_2^i.
\end{aligned}$$

Equating coefficients gives (2.31a). In a similar manner, by differentiating (2.8) with respect to $\overline{w}$, w and $\overline{z}$ we may obtain (2.31b), (2.31c) and (2.31d) respectively. ∎

Lemma 2.16. *Any of the first partial derivatives map the outer functions of degree $-m-2$ to (multiples of) outer functions of degree $-m-3$. Specifically, for $m \ge 0$ and $0 \le j \le m$,*

$$\frac{\partial}{\partial z} o_{i,j}^{m}(z,w) = -(m+1-i)\ o_{i,j}^{m+1}(z,w), \qquad 0 \le i \le m+1, \tag{2.32a}$$
$$\frac{\partial}{\partial \overline{w}} o_{i,j}^{m}(z,w) = (m+1-i)\ o_{i,j+1}^{m+1}(z,w), \qquad 0 \le i \le m+1, \tag{2.32b}$$

$$\frac{\partial}{\partial w} o_{i,j}^m(z,w) = -(i+1)\ o_{i+1,j}^{m+1}(z,w), \qquad -1 \le i \le m, \tag{2.32c}$$

$$\frac{\partial}{\partial \overline{z}} o_{i,j}^m(z,w) = -(i+1)\ o_{i+1,j+1}^{m+1}(z,w), \qquad -1 \le i \le m. \tag{2.32d}$$

Proof of (2.32a). The proof is by induction on m.

Induction basis. Since

$$o_{0,0}^0(z,w) = (z\overline{z} + w\overline{w})^{-1},$$

an Equation (2.27) shows that

$$O_1(z,w) = \begin{bmatrix} o_{0,0}^1(z,w) & o_{0,1}^1(z,w) \\ o_{1,0}^1(z,w) & o_{1,1}^1(z,w) \end{bmatrix} = |x|^{-4} \begin{bmatrix} \overline{z} & -w \\ \overline{w} & z \end{bmatrix}.$$

Direct differentiation of $o_{0,0}^0$ by z, $\overline{w}$, w, and $\overline{z}$ gives $o_{0,0}^1$, $o_{0,1}^1$, $o_{1,0}^1$ and $o_{1,1}^1$, respectivily, and hence Equations (2.32) are true for $m = 0$.

Induction step. Assume that (2.32a) is true for $0 \le m \le M$. By the first part of (2.27),

$$o_{i,j}^{M+1}(z,w) = \frac{\overline{z}}{|x|^2} o_{i,j}^M(z,w) + \frac{\overline{w}}{|x|^2} o_{i-1,j}^M(z,w),$$

for $0 \le i \le M+1$ and $0 \le j \le M$. Using the inductive hypothesis to differentiate the outer functions,

$$\begin{aligned}
\frac{\partial}{\partial z} o_{i,j}^{M+1}(x) =& \frac{\overline{z}(-\overline{z})}{|x|^4} o_{i,j}^M(x) + \frac{\overline{z}}{|x|^2}\frac{\partial}{\partial z} o_{i,j}^M(x) + \frac{\overline{w}(-\overline{z})}{|x|^4} o_{i-1,j}^M(x) + \frac{\overline{w}}{|x|^2}\frac{\partial}{\partial z} o_{i-1,j}^M(x) \\
=& \frac{-\overline{z}}{|x|^2}\left(\frac{\overline{z}}{|x|^2} o_{i,j}^M(x) + \frac{\overline{w}}{|x|^2} o_{i-1,j}^M(x)\right) - (M+1-i)\frac{\overline{z}}{|x|^2} o_{i,j}^{M+1}(x) \\
& - (M+2-i)\frac{\overline{w}}{|x|^2} o_{i-1,j}^{M+1}(x).
\end{aligned}$$

Two applications of the first part of (2.27) give the result.

However, if $j = M+1$ the above does not hold. By the second part of (2.27),

$$o_{i,j}^{M+1}(z,w) = \frac{-w}{|x|^2} o_{i,j-1}^M(z,w) + \frac{z}{|x|^2} o_{i-1,j-1}^M(z,w),$$

for $0 \le i \le M+1$ and $1 \le j \le M+1$. Once again the inductive hypothesis will be used to differentiate this expression. This gives

$$\begin{aligned}
\frac{\partial}{\partial z} o_{i,j}^{M+1}(x) =& \frac{(-w)(-\overline{z})}{|x|^4} o_{i,j-1}^M(x) + \frac{(-w)}{|x|^2}\frac{\partial}{\partial z} o_{i,j-1}^M(x) \\
& + \frac{z(-\overline{z})}{|x|^4} o_{i-1,j-1}^M(x) + \frac{z}{|x|^2}\frac{\partial}{\partial z} o_{i-1,j-1}^M(x) + \frac{1}{|x|^2} o_{i-1,j-1}^M(x) \\
=& \frac{(-w)(-\overline{z})}{|x|^4} o_{i,j-1}^M(x) - (M+1-i)\frac{(-w)}{|x|^2} o_{i,j-1}^{M+1}(x) \\
& + \frac{z(-\overline{z})}{|x|^4} o_{i-1,j-1}^M(x) - (M+2-i)\frac{z}{|x|^2} o_{i-1,j-1}^{M+1}(x) + \frac{z\overline{z} + w\overline{w}}{|x|^4} o_{i-1,j-1}^M(x)
\end{aligned}$$

$$
\begin{aligned}
&= -(M+1-i)\frac{(-w)}{|x|^2}o_{i,j-1}^{M+1}(x) - (M+2-i)\frac{z}{|x|^2}o_{i-1,j-1}^{M+1}(x) \\
&\quad - \frac{(-w)}{|x|^2}\left(\frac{\overline{z}}{|x|^2}o_{i,j-1}^{M}(x) + \frac{\overline{w}}{|x|^2}o_{i-1,j-1}^{M}(x)\right).
\end{aligned}
$$

By the first part of (2.27), the term in the brackets is equal to $o_{i,j-1}^{M+1}(z,w)$ and thus the second part of (2.27) gives the required result. This proves (2.32a) by induction on m. The other three relations in (2.32) may be proven in a similar way. ∎

Given the above expressions for the derivatives of the inner and outer functions it is easy to show that they are harmonic functions. Since (2.30) may be used for the Laplacian, it follows that

$$
\frac{1}{4}\Delta o_{i,j}^m = \big(-(m+1-i)\big)\big(-(i+1)\big)o_{i+1,j+1}^{m+2} + (m+1-i)\big(-(i+1)\big)o_{i+1,j+1}^{m+2} = 0, \quad (2.33)
$$

and

$$
\frac{1}{4}\Delta t_{i,j}^{m+2} = j(m+2-j)t_{i-1,j-1}^{m} - j(m+2-j)t_{i-1,j-1}^{m} = 0. \quad (2.34)
$$

Each of (2.33) and (2.34) can be inferred from the other as indicated in Remark 2.9.

We now wish to show that $|\cdot|^{2\ell}t_{i,j}^m$ and $|\cdot|^{2\ell}o_{i,j}^m$ are polyharmonic functions. To do this we will need two general results.

Lemma 2.17 (Euler's Homogeneous Function Theorem). *If $f:\mathbb{R}^d\to\mathbb{R}$ is an homogeneous function of order m,* i.e., *$f(\alpha x)=\alpha^m f(x)$, then*

$$
x\cdot(\nabla f)(x) = mf(x). \quad (2.35)
$$

Proof. Differentiating $\alpha^m f(x) = f(\alpha x)$ with respect to α gives

$$
m\alpha^{m-1}f(x) = \sum_{i=1}^{d} x_i\frac{\partial f}{\partial x_i}(\alpha x) = x\cdot(\nabla f)(\alpha x).
$$

Setting $\alpha = 1$ gives the result. ∎

Lemma 2.18. *Let $|\cdot|$ be the 2-norm on $\mathbb{R}^d$. If $f:\mathbb{R}^d\to\mathbb{R}$ is a harmonic function that is homogeneous of degree m then*

$$
\Delta(|\cdot|^{2\ell}f) = 4\ell\left(\frac{d}{2}+\ell+m-1\right)|\cdot|^{2(\ell-1)}f.
$$

Furthermore, for d even, $|\cdot|^{2\ell}f$ is polyharmonic of order

$$
\begin{cases}
\ell+1, & \text{for } \ell\ge 0 \text{ and } m>-d/2 \text{ or } m<1-\ell-d/2,\\
\ell+m+d/2, & \text{for } \ell<0 \text{ and } m\ge 1-\ell-d/2.
\end{cases}
$$

Proof. From the product rule for the Laplacian,

$$\Delta(fg) = (\Delta f)g + 2(\nabla f)\cdot(\nabla g) + f(\Delta g),$$

and the Euler relation (2.35), we obtain

$$\begin{aligned}
\Delta\big(|x|^{2\ell}f(x)\big) &= |x|^{2\ell}\Delta f(x) + 2\nabla|x|^{2\ell}\cdot\nabla f(x) + f(x)\Delta|x|^{2\ell} \\
&= 4\ell|x|^{2(\ell-1)}x\cdot\nabla f(x) + f(x)\left(2d\ell|x|^{2(\ell-1)} + 4\ell(\ell-1)|x|^{2(\ell-2)}\sum_{j=1}^{d}x_j^2\right) \\
&= 4\ell m|x|^{2(\ell-1)}f(x) + \big(2d\ell + 4\ell(\ell-1)\big)|x|^{2(\ell-1)}f(x) \\
&= 4\ell(d/2+\ell+m-1)|x|^{2(\ell-1)}f(x),
\end{aligned}$$

which shows the first part of the lemma. Observing the number of applications of Δ that are required to reduce one of the multipliers to zero gives the second part. ∎

Since the outer functions $o_{i,j}^m : \mathbb{R}^4 \to \mathbb{R}$ are harmonic and homogeneous of degree $-m-2$ and the inner functions $t_{i,j}^m : \mathbb{R}^4 \to \mathbb{R}$ are harmonic and homogeneous of degree m, Lemma 2.18 implies the following result.

Corollary 2.19. *The functions $|\cdot|^{2\ell}t_{i,j}^m$ and $|\cdot|^{2\ell}o_{i,j}^m$ are polyharmonic of order $\ell+1$ for $\ell = 0,1,\ldots,m$. Furthermore,*

$$\begin{aligned}
\Delta(|\cdot|^{2\ell}o_{i,j}^m) &= 4\ell(\ell-m-1)|\cdot|^{2(\ell-1)}o_{i,j}^m, \\
\Delta(|\cdot|^{2\ell}t_{i,j}^m) &= 4\ell(\ell+m+1)|\cdot|^{2(\ell-1)}t_{i,j}^m.
\end{aligned}$$

All of the outer functions $o_{i,j}^m(x)$ may be written as (multiples of) derivatives of $1/|x|^2$ as in the following.

Lemma 2.20. *Define*

$$R_{i,j}^m = \begin{cases} \dfrac{(-1)^{m-j+i}}{i!(m-i)!}\dfrac{\partial^m}{\partial\overline{w}^{j-i}\partial z^{m-j}\partial\overline{z}^i}, & i \le j, \\[2ex] \dfrac{(-1)^m}{i!(m-i)!}\dfrac{\partial^m}{\partial w^{i-j}\partial z^{m-i}\partial\overline{z}^j}, & i \ge j, \end{cases} \tag{2.36}$$

$0 \le i,j \le m$, $m \in \mathbb{N}_0$. *Then for $x \in \mathbb{R}^4\backslash\{0\}$,*

$$o_{i,j}^m(x) = R_{i,j}^m\frac{1}{|x|^2}. \tag{2.37}$$

Furthermore, for harmonic functions f,

$$R_{i,j}^m R_{i',j'}^{m'}\, f = e(m,m',i)_{i',i'}R_{i+i',j+j'}^{m+m'}\, f, \tag{2.38}$$

where

$$e(m, m', i)_{i',i'} = \binom{i+i'}{i}\binom{m+m'-(i+i')}{m-i}. \tag{2.39}$$

Proof. The definition of the outer functions in terms of the inner functions, (2.16), and the fact that $t^0_{0,0} = 1$, imply

$$o^0_{0,0}(x) = \frac{1}{|x|^2}.$$

Then (2.37) follows by repeated use of Lemma 2.16.

Consider now (2.38). If $i \le j$ and $i' \le j'$ *or* $i \ge j$ and $i' \ge j'$, then (2.38) follows easily from (2.36). Turning to the slightly more difficult *mixed case* assume, without loss of generality, $i \le j$ and $i' \ge j'$. Using Definition (2.36),

$$R^m_{i,j} R^{m'}_{i',j'} f = \frac{(-1)^{m-j+i+m'}}{i!(m-i)!i'!(m'-i')!} \frac{\partial^{m+m'}}{\partial \overline{w}^{j-i} \partial w^{i'-j'} \partial z^{m-j+m'-i'} \partial \overline{z}^{i+j'}} f. \tag{2.40}$$

Let g be harmonic. Then from the form (2.30) for the Laplacian

$$\frac{\partial^2}{\partial z \partial \overline{z}} g = -\frac{\partial^2}{\partial w \partial \overline{w}} g.$$

Hence, if $i' - j' \ge j - i$, then

$$\frac{\partial^{i'-j'+j-i}}{\partial \overline{w}^{j-i} \partial w^{i'-j'}} f = (-1)^{j-i} \frac{\partial^{i'-j'+j-i}}{\partial w^{i+i'-(j+j')} \partial z^{j-i} \partial \overline{z}^{j-i}} f.$$

Substituting this into (2.40) gives

$$\begin{aligned} R^m_{i,j} R^{m'}_{i',j'} f &= \frac{(-1)^{m+m'}}{i!(m-i)!i'!(m'-i')!} \frac{\partial^{m+m'}}{\partial w^{i+i'-(j+j')} \partial z^{m+m'-(i+i')} \partial \overline{z}^{j+j'}} f \\ &= e(m, m', i)_{i',i'} R^{m+m'}_{i+i',j+j'} f, \end{aligned}$$

since $i' - j' \ge j - i$ implies $i + i' \ge j + j'$. The case when $i' - j' \le j - i$ is similar. ∎

Definition 2.21. If V is an n-dimensional vector space and $X = \{x_1, \ldots, x_n\}$ is a basis for V, then there is a uniquely determined basis $X' = \{y_1, \ldots, y_n\}$ for V' with the property that

$$[x_i, y_j] = \delta_{i,j}, \qquad 1 \le i, j \le n,$$

where $[\cdot, \cdot]$ denotes the action of an element of V' on an element of V. This basis is called the (bi-orthogonal) *dual* of the basis X.

Motivated by the operators $R^m_{i,j}$ we are led to define operators

$$L^m_{i,j} = \begin{cases} \dfrac{1}{j!(m-j)!} \dfrac{\partial^m}{\partial w^{j-i} \partial \overline{z}^i \partial z^{m-j}} & i \le j, \\[2ex] \dfrac{(-1)^{i-j}}{j!(m-j)!} \dfrac{\partial^m}{\partial \overline{w}^{i-j} \partial \overline{z}^j \partial z^{m-i}} & i \ge j. \end{cases} \tag{2.41}$$

These operators can be used to specify a dual basis for the inner functions. Repeated use of Lemma 2.15 leads to the following lemma.

Lemma 2.22 (Dual basis for the inner functions). *The operators $L^m_{i,j}$ satisfy*

$$L^m_{i,j} t^m_{i,j} = t^0_{0,0},$$

and also have the more general property

$$L^{m'}_{i',j'} t^m_{i,j} = \binom{j}{j'} \binom{m-j}{m'-j'} t^{m-m'}_{i-i',j-j'}.$$

Thus the functionals

$$\lambda^m_{i,j}(f) = \big(L^m_{i,j} f\big)(0), \tag{2.42}$$

$0 \le i, j \le m$, $0 \le m \le p$, *form a dual basis for* $\left\{t^m_{i,j},\ 0 \le i, j \le m,\ 0 \le m \le p\right\}$.

2.3 Expansions of polyharmonic basic functions

In this section we develop far field expansions of the functions $|\cdot - x_<|^{-2}$, and $|\cdot - x_<|^{2n} \ln(|\cdot - x_<|)$, $n = 0, 1, \ldots$. These functions lead to polyharmonic splines of order 1 and $n+2$ respectively. We also find bounds on the error in approximating the associated polyharmonic radial basis functions by truncating these far field expansions. We will truncate by dropping all terms of sufficiently negative homogeneity.

We will find it useful in this section to make use of the cosine formula in the form

$$\begin{aligned}
|x - x_<|^2 &= |x|^2 + |x_<|^2 - 2\langle x, x_< \rangle \\
&= |x|^2 + |x_<|^2 - |x|^2 \operatorname{Tr}(x^{-1} x_<) \\
&= |x|^2 + |x_<|^2 - |x|^2 \chi_1(x^{-1} x_<) \qquad (2.43a) \\
&= |x|^2 + |x_<|^2 - \chi_1(x^* x_<) \\
&= |x|^2 + |x_<|^2 - \chi_1(x_<{}^* x), \qquad (2.43b)
\end{aligned}$$

where the second line follows from the first when we recall the defintion of the inner product from Equation (2.5).

2.3.1 Expansion of the potential

We start with a far field expansion of the potential function $|x - x_<|^{-2}$. This is an expansion in the character functions, which are products of powers of $|x_<|/|x|$ and the appropriate Gengenbauer polynomials—the Chebyshev polynomials of the second kind.

Lemma 2.23. *For $x, x_< \in \mathbb{R}^4$ with $|x_<| < |x|$,*

$$\frac{1}{|x - x_<|^2} = \frac{1}{|x|^2} \sum_{m=0}^{\infty} \chi_m(x^{-1}x_<). \tag{2.44}$$

Proof. The result is trivially true when $x_< = 0$. Hence in what follows we assume that $0 < |x_<| < |x|$. Then from the definition of the character function, (2.13),

$$\chi_m(x^{-1}x_<) = \left(\frac{|x_<|}{|x|}\right)^m \frac{\sin(m+1)\theta}{\sin\theta},$$

where θ is the angle between x and $x_<$. As is well known,

$$\left|\frac{\sin(m+1)\theta}{\sin\theta}\right| \leq m+1, \qquad \forall\, \theta \in \mathbb{R}.$$

Therefore the series on the right of (2.44) converges absolutely.

We will prove the lemma by showing the product of the right hand side of (2.44) with $|x - x_<|^2$ is 1. Let $y = x^{-1}x_<$, then

$$\begin{aligned}
|x - x_<|^2 \frac{1}{|x|^2} \sum_{m=0}^{\infty} \chi_m(y) &= \left\{|x|^2 + |x_<|^2 - |x|^2\chi_1(y)\right\} \frac{1}{|x|^2} \sum_{m=0}^{\infty} \chi_m(y) \\
&= \left\{1 + |y|^2 - \chi_1(y)\right\} \sum_{m=0}^{\infty} \chi_m(y) \\
&= \sum_{m=0}^{\infty} \left\{\chi_m(y) + |y|^2\chi_m(y) - \chi_1(y)\chi_m(y)\right\} \\
&= \sum_{m=0}^{\infty} \left\{\chi_m(y) + |y|^2\chi_m(y) - \chi_{m+1}(y) - |y|^2\chi_{m-1}(y)\right\} \\
&= \chi_0(y) - |y|^2\chi_{-1}(y) \\
&= 1,
\end{aligned}$$

where we have used Lemma 2.8 to expand the product $\chi_1(y)\chi_m(y)$ and the fact that $\chi_{-1} = 0$. Note that the telescoping argument is valid because $\chi_m \to 0$ as $m \to \infty$. ∎

We now wish to obtain a bound on the error in approximating $\Phi_{x_<}(\cdot) = 1/|\cdot - x_<|^2$ by the truncated series

$$g_p(x) = \frac{1}{|x|^2} \sum_{m=0}^{p} \chi_m(x^{-1}x_<).$$

From the explicit formula for the character function, (2.13), $|\chi_m(y)| \leq (m+1)|y|^m$ for all $y \in \mathbb{H}^1$ and therefore the error in approximating $\Phi_{x_<}(\cdot)$ by g_p is bounded by

$$\left|\Phi_{x_<}(x) - g_p(x)\right| \leq \frac{1}{|x|^2} \sum_{m=p+1}^{\infty} \left|\chi_m(x^{-1}x_<)\right|$$

$$\begin{aligned}
&\leq \frac{1}{|x|^2} \sum_{m=p+1}^{\infty} (m+1)|y|^m \\
&= \frac{|y|^{p+1}}{|x|^2} \sum_{m=0}^{\infty} \big((p+1)+(m+1)\big)|y|^m \\
&= \frac{|y|^{p+1}}{|x|^2} \left\{ (p+1) \sum_{m=0}^{\infty} |y|^m + \sum_{m=0}^{\infty} (m+1)|y|^m \right\}.
\end{aligned}$$

where $y = x^{-1}x_<$. If $|x_<| < |x|$ then $|y| < 1$ and the well known identities

$$\sum_{m=0}^{\infty} h^m = \frac{1}{1-h} \qquad \text{and} \qquad \sum_{m=0}^{\infty} (m+1)h^m = \frac{1}{(1-h)^2},$$

for $|h| < 1$, may be applied. This gives

$$\left|\Phi_{x_<}(x) - g_p(x)\right| \leq \frac{|y|^{p+1}}{|x|^2} \left\{ \frac{p+1}{1-|y|} + \frac{1}{(1-|y|)^2} \right\} \tag{2.45}$$

Denote the bound on the right of (2.45) by $e_p(|y|)$. For $|x|$ fixed, since each term on the right in (2.45) is obviously strictly increasing in $|y|$ for $0 < |y| < 1$, so is $e_p(|y|)$. Considering now the sum

$$s(x) = \sum_{k=1}^{N} \frac{d_k}{|x-x_k|^2},$$

we apply the bound above to each term and sum. The monotonicity of the bound enables us to estimate $e_p(|x_k|/|x|)$ by $e_p(d)$ where

$$d = \max_{1 \leq k \leq N} \frac{|x_k|}{|x|}.$$

In combination with Lemma 2.23 and (2.18), this gives

Theorem 2.24. *Suppose $x_k \in \mathbb{R}^4$, $|x_k| \leq r$ and $d_k \in \mathbb{R}$ for each $1 \leq k \leq N$. Let*

$$s(x) = \sum_{k=1}^{N} \frac{d_k}{|x-x_k|^2},$$

and let C_m be the $(m+1) \times (m+1)$ matrix

$$[C^m_{i,j}] = \sum_{k=1}^{N} d_k \big(T_m(x_k)\big).$$

For $p \in \mathbb{N}_0$, set

$$g_p(x) = \sum_{m=0}^{p} \sum_{i,j=0}^{m} C^m_{j,i}\, o^m_{i,j}(x) = \sum_{m=0}^{p} \operatorname{Tr}\left(C_m O_m(x)\right), \tag{2.46}$$

$x \in \mathbb{R}^4 \backslash \{0\}$. Then for all x with $|x| > r$

$$|s(x) - g_p(x)| \leq \frac{M}{r^2} \left(\frac{p+1}{1-1/c} + \frac{1}{(1-1/c)^2} \right) \left(\frac{1}{c}\right)^{p+3}$$

where $M = \sum_{k=1}^{N} |d_k|$ and $c = |x|/r$.

2.3.2 Expansion of a polyharmonic function

In this subsection our aim is to develop far field expansions for the functions $|\cdot - x_<|^{2n} \ln|\cdot - x_<|$, which are polyharmonic of order $n+2$. This will be done by induction on n with the biharmonic case $\ln|\cdot - x_<|$ being used as the induction basis.

The polyharmonic functions f of order $n+2$ that occur will be written in the form

$$f = f_0 + |\cdot|^2 f_1 + \cdots + |\cdot|^{2n+2} f_{n+1}$$

where $f_0, \ldots, f_{n+1}$ are harmonic. In this sum, $|\cdot|^{2j} f_j$ is actually a $(j+1)$-harmonic term. As a consequence the terms of a specified homogeneous order k in our expansions will no longer involve a single χ_m as in the harmonic case of Lemma 2.23. Rather they will be a weighted sum of $\chi_m(y)$, $|y|^2\chi_{m-2}(y), \ldots, |y|^{2n+2}\chi_{m-2n-2}(y)$, consistent with the polyharmonicity orders of Lemma 2.18. For this reason we need to know how pairs of character functions combine.

Lemma 2.25. *For $m \in \mathbb{N}$, $m \geq 2$ and $|x_<| < |x|$*

$$\chi_m(x^{-1}x_<) - \frac{|x_<|^2}{|x|^2}\chi_{m-2}(x^{-1}x_<) = 2\frac{|x_<|^m}{|x|^m}\cos(m\theta), \tag{2.47}$$

where θ is the angle between x and $x_<$.

Proof. The Lemma is trivially true when $x_< = 0$. For $0 < |x_<| < |x|$, the explicit formula for the character function implies

$$\begin{aligned}
\chi_m(x^{-1}x_<) - \frac{|x_<|^2}{|x|^2}\chi_{m-2}(x^{-1}x_<) &= \frac{|x_<|^m}{|x|^m}\left(\frac{\sin(m+1)\theta}{\sin\theta} - \frac{\sin(m-1)\theta}{\sin\theta}\right) \\
&= \frac{|x_<|^m}{|x|^m}\left(\frac{\sin(m+1)\theta - \sin(m-1)\theta}{\sin\theta}\right) \\
&= 2\frac{|x_<|^m}{|x|^m}\frac{\sin(\theta)\cos(m\theta)}{\sin\theta} \\
&= 2\frac{|x_<|^m}{|x|^m}\cos(m\theta).
\end{aligned}$$

∎

We now proceed to obtain a far field expansion for the biharmonic function $\ln|\cdot - x_<|$. While this expansion is useful in and of itself, we will also use it as the induction basis for the expansion of the more general function $|\cdot - x_<|^{2n}\ln|\cdot - x_<|$, which is polyharmonic of order $n+2$.

Lemma 2.26. *For $x, x_< \in \mathbb{R}^4$ and $|x_<| < |x|$,*

$$\ln|x - x_<|^2 = \ln|x|^2 - \sum_{m=1}^{\infty}\frac{1}{m}\left\{\chi_m(x^{-1}x_<) - \frac{|x_<|^2}{|x|^2}\chi_{m-2}(x^{-1}x_<)\right\}. \tag{2.48}$$

Proof. The case $x_< = 0$ is trivially true. Hence we assume $0 < |x_<| < |x|$.

We will use $\ln(\cdot)$ for the real logarithm and $\log(\cdot)$ for the principal branch of the complex logarithm. Thus $\Re(\log z) = \ln|z|$ away from the branch cut. We will represent $x = (x_1, x_2, x_3, x_4)^T \in \mathbb{R}^4$ by $x = [x_1 + \mathrm{i}x_2, x_3 + \mathrm{i}x_4] \in \mathbb{C}^2$ and similarly for $x_<$. There exists a rotation R_1 such that

$$R_1 x = [|x|, 0].$$

By the argument that precedes the introduction of the character function χ_m in (2.13), there is a rotation R_2 that fixes the north pole $[1, 0]$ and rotates $R_1 x_<$ to the direction $[e^{\mathrm{i}\theta}, 0]$ where θ is the angle between $R_1 x_<$ and the north pole. Since rotations preserve angles θ is also the angle between x and $x_<$, $R = R_2 R_1$ is a rotation such that

$$Rx = [|x|, 0], \qquad Rx_< = [|x_<|e^{\mathrm{i}\theta}, 0].$$

Thus

$$|x - x_<| = |R(x - x_<)| = |Rx - Rx_<| = \left|[|x| - |x_<|e^{\mathrm{i}\theta}, 0]\right| = |x|\left|1 - \frac{|x_<|}{|x|}e^{\mathrm{i}\theta}\right| = |x|\left|1 - |y|e^{\mathrm{i}\theta}\right|,$$

where $y = x^{-1}x_<$, and

$$\begin{aligned}\ln|x - x_<| &= \Re\left\{\log\left(|x|\left(1 - |y|e^{\mathrm{i}\theta}\right)\right)\right\} = \Re\left\{\log(|x|) + \log\left(1 - |y|e^{\mathrm{i}\theta}\right)\right\} \\ &= \ln|x| + \Re\left\{\log\left(1 - |y|e^{\mathrm{i}\theta}\right)\right\}.\end{aligned}$$

Now

$$\log\left(1 - |y|e^{\mathrm{i}\theta}\right) = -\sum_{m=1}^{\infty}\frac{1}{m}\left(|y|e^{\mathrm{i}\theta}\right)^m = -\sum_{m=1}^{\infty}\frac{1}{m}|y|^m e^{\mathrm{i}m\theta}.$$

Taking the real part of this expression,

$$\begin{aligned}\Re\left\{\log\left(1 - \frac{|x_<|}{|x|}e^{\mathrm{i}\theta}\right)\right\} &= -\sum_{m=1}^{\infty}\frac{1}{m}\frac{|x_<|^m}{|x|^m}\cos(m\theta) \\ &= -\frac{1}{2}\sum_{m=1}^{\infty}\frac{1}{m}\left\{\chi_m(x^{-1}x_<) - \frac{|x_<|^2}{|x|^2}\chi_{m-2}(x^{-1}x_<)\right\},\end{aligned}$$

where we have used Lemma 2.25 to express $(|x_<|^m/|x|^m)\cos(m\theta)$ in terms of the character functions χ_m. ∎

We now wish to obtain an expansion for the polyharmonic function

$$|x - x_<|^{2n}\ln|x - x_<|^2.$$

To simplify this procedure, we observe that

$$|x - x_<|^{2n} \ln |x - x_<|^2 = |x - x_<|^{2n} \ln |x|^2 + |x|^{2n} |I - x^{-1}x_<|^{2n} \ln |I - x^{-1}x_<|^2, \tag{2.49}$$

where I is the 2×2 identity in $\mathbb{H}_0^1$ or the element $[1, 0]$ in $\mathbb{C}^2$. This splits the function into a term containing the logarithmic singularity and a term amenable to "Laurent" expansion. We shall handle the two parts of the right hand side of (2.49) separately.

The coefficient of $\ln |x|$ in the expansion

We first consider the polynomial that multiplies the $\ln |x|$ term in (2.49). We will give an expression for this polynomial in terms of the inner functions and discuss some symmetry properties.

Lemma 2.27. *For $x, x_< \in \mathbb{R}^4$,*

$$\begin{aligned} |x - x_<|^{2n} &= \sum_{m=0}^{n} |x|^{2m} \sum_{\ell=0}^{n-m} b^n_{m,\ell} |x_<|^{2\ell} \chi_{n-m-\ell}(x_<{}^* x) \\ &= \sum_{m=0}^{n} |x|^{2m} \sum_{\ell=0}^{n-m} \sum_{i,j=0}^{n-m-\ell} D^{m,\ell}_{j,i}(x_<) t^{n-m-\ell}_{i,j}(x) \end{aligned} \tag{2.50}$$

where the coefficients $b^n_{m,\ell}$ are given recursively by

$$b^{n+1}_{m,\ell} = b^n_{m-1,\ell} + b^n_{m,\ell-1} - b^n_{m,\ell} - b^n_{m-1,\ell-1} \tag{2.51a}$$

along with the initial conditions

$$b^0_{m,\ell} = \begin{cases} 1 & m = \ell = 0, \\ 0 & \text{otherwise}, \end{cases} \tag{2.51b}$$

$$b^n_{m,\ell} = 0 \qquad \text{if } m + \ell > n, \text{ or } m, \ell \notin [0, n], \tag{2.51c}$$

and the coefficients $D^{m,\ell}_{j,i}(x_<)$ are given by

$$D^{m,\ell}_{j,i}(x_<) = b^n_{m,\ell} |x_<|^{2\ell} t^{n-m-\ell}_{j,i}(x_<{}^*). \tag{2.52}$$

Proof. The proof of the first equality in Equation (2.50) is by induction on n

Induction basis. Since $\chi_0 = 1$, in the case $n = 0$ the first equality in Equation (2.50) is just $1 = b^0_{0,0}$, and hence the lemma is true for $n = 0$.

Induction step. Assume the lemma is true for $n = N$. Using the cosine formula in the form of Equation (2.43b) and the product rule (2.14) for character functions

$$\begin{aligned}
|x - x_<|^{2(N+1)} &= |x - x_<|^2 |x - x_<|^{2N} \\
&= \left(|x|^2 + |x_<|^2 - \chi_1 x_<{}^* x\right)\left(\sum_{m=0}^{N}\sum_{\ell=0}^{N-m} b^N_{m,\ell}|x|^{2m}|x_<|^{2\ell}\chi_{N-m-\ell}(x_<{}^*x)\right) \\
&= \sum_{m=0}^{N}\sum_{\ell=0}^{N-m} b^N_{m,\ell}|x|^{2(m+1)}|x_<|^{2\ell}\chi_{N-m-\ell}(x_<{}^*x) \\
&\qquad + b^N_{m,\ell}|x|^{2m}|x_<|^{2(\ell+1)}\chi_{N-m-\ell}(x_<{}^*x) \\
&\qquad - b^N_{m,\ell}|x|^{2m}|x_<|^{2\ell}\chi_{N-m-\ell+1}(x_<{}^*x) \\
&\qquad - b^N_{m,\ell}|x|^{2(m+1)}|x_<|^{2(\ell+1)}\chi_{N-m-\ell-1}(x_<{}^*x) \\
&= \sum_{m=0}^{N+1}\sum_{\ell=0}^{N+1-m} b^{N+1}_{m,\ell}|x|^{2m}|x_<|^{2\ell}\chi_{N+1-m-\ell}(x_<{}^*x),
\end{aligned}$$

where

$$b^{N+1}_{m,\ell} = b^N_{m-1,\ell} + b^N_{m,\ell-1} - b^N_{m,\ell} - b^N_{m-1,\ell-1}$$

as long as $b^N_{m,\ell}$ is taken to be zero when $m + \ell > N$ or $m, \ell \notin [0, N]$. This proves the first equality in Equation (2.50) by induction. The second equality in Equation (2.50) follows immediately from the first by substituting expression (2.19) for $\chi_{n-m-\ell}(x_<{}^*x)$. ∎

Remark 2.28. Since $b^n_{m,\ell}$ is real and

$$t^{n-m-\ell}_{j,i}(x_<{}^*) = (-1)^{i+j}\overline{t^{n-m-\ell}_{n-m-\ell-j,n-m-\ell-i}(x_<{}^*)}$$

by symmetry (2.22) we see that

$$D^{m,\ell}_{j,i}(x_<) = (-1)^{i+j}\overline{D^{m,\ell}_{n-m-\ell-j,n-m-\ell-i}(x_<)}$$

for all $0 \le m \le n$, $0 \le \ell \le n - m$, $0 \le i, j \le n - m - \ell$. Provided the weights d_k are real, this symmetry is inherited by the coefficients of polynomials

$$q(x) = \sum_{k=0}^{n} d_k|x - x_k|^{2n} = \sum_{m=0}^{n}|x|^{2m}\sum_{\ell=0}^{n-m}\sum_{i,j=0}^{n-m-\ell}\widetilde{D}^{m,\ell}_{j,i}t^{n-m-\ell}_{i,j}(x) \tag{2.53}$$

occurring as the coefficient of the $\ln|x|^2$ in the truncated far field expansion of Theorem 2.32 to come.

One use of this property would be to recast the polynomial $q(x)$ as the weighted sum of approximately half as many $t^{n-m-\ell}_{i,j}(x)$'s, thereby reducing the operation count for approximate evaluation.

The non-logarithmic part in the expansion

We now consider the infinite or far field part of (2.49). We find an explicit form for the expansion and give bounds on the error in approximation by truncation for this series.

Lemma 2.29. *For $n \in \mathbb{N}_0$ and $y \in \mathbb{R}^4$, $|y| < 1$,*

$$|I - y|^{2n} \ln |I - y|^2 = \sum_{\ell=0}^{n+1} \sum_{m=\max\{1,2\ell\}}^{\infty} c_{m,\ell}^n |y|^{2\ell} \chi_{m-2\ell}(y) \tag{2.54}$$

where the series (2.54) *converges absolutely. The coefficients $c_{m,\ell}^n$ are given by the formulae*

$$c_{m,\ell}^0 = \begin{cases} -1/m, & \ell = 0,\ m \geq 1, \\ 1/m, & \ell = 1,\ m - 2\ell \geq 0, \\ 0, & \text{otherwise}, \end{cases} \tag{2.55a}$$

and the recurrence

$$c_{m,\ell}^{n+1} = c_{m,\ell}^n - c_{m-1,\ell}^n - c_{m-1,\ell-1}^n + c_{m-2,\ell-1}^n. \tag{2.55b}$$

Proof. The proof is by induction on n.

Induction Basis: The case $n = 0$ of formulae (2.54) and (2.55a) is contained in Lemma 2.26.

Induction Step: Assume (2.54) has been established for $n = K$. Then using the cosine formula (2.43a) and the product formula (2.14) we find

$$\begin{aligned} &|I - y|^{2K+2} \ln |I - y|^2 \\ &= \sum_{\ell=0}^{K+1} \sum_{m=\max\{1,2\ell\}}^{\infty} c_{m,\ell}^K |y|^{2\ell} \Big\{ \chi_{m-2\ell}(y) - \chi_{m+1-2\ell}(y) - |y|^2 \chi_{m-1-2\ell}(y) + |y|^2 \chi_{m-2\ell}(y) \Big\}. \end{aligned}$$

Rearranging by collecting terms of the same homogeneity we find a series

$$|I - y|^{2K+2} \ln |I - y|^2 = \sum_{\ell=0}^{K+2} \sum_{m=\max\{1,2\ell\}}^{\infty} c_{m,\ell}^{K+1} |y|^{2\ell} \chi_{m-2\ell}(y)$$

converging absolutely for $|y| < 1$ and with coefficients given by (2.55b). Thus (2.54) holds for $n = K + 1$. The result follows by induction. ∎

Lemma 2.30. *For $m > 2n \geq 0$ and $0 \leq \ell \leq n + 1$, the coefficients $c_{m,\ell}^n$ defined recursively in Lemma 2.29 have the explicit form*

$$c_{m,\ell}^n = (-1)^{n+\ell+1} n! \binom{n+1}{\ell} \Bigg/ \prod_{\substack{k=m-\ell-n \\ k \neq m-2\ell+1}}^{m-\ell+1} k. \tag{2.56}$$

Proof. The proof is by induction on n.

Induction Basis: The formula for $n = 0$ is formula (2.55a) of Lemma 2.29.

Induction Step: Assume that the formula holds for $n = K$ and $m > 2K$. Then for $n = K+1$

$$c^{K+1}_{m,\ell} = c^K_{m,\ell} - c^K_{m-1,\ell} - c^K_{m-1,\ell-1} + c^K_{m-2,\ell-1}.$$

Assume now that $m > 2(K+1)$ and $1 \le \ell \le K+1$. Using the induction hypothesis,

$$\begin{aligned}
c^{K+1}_{m,\ell} =&(-1)^{K+1+\ell} K! \Big\{ \binom{K+1}{\ell} \Big\{ \big(m-\ell-(K+1)\big)(m-2\ell+1) \\
&- (m-\ell+1)(m-2\ell) \Big\} + \binom{K+1}{\ell-1} \Big\{ (m-\ell-K-1)(m-2\ell+2) \\
&- (m-\ell+1)(m-2\ell+1) \Big\} \Big\} \Big/ \prod_{k=m-\ell-(K+1)}^{m-\ell+1} k \\
=&(-1)^{K+1+\ell} \frac{K!(K+1)!}{\ell!(K+2-\ell)!} \Big\{ (K+2-\ell) \Big\{ (m-\ell-K-1)(m-2\ell+1) \\
&- (m-\ell-1)(m-2\ell) \Big\} + \ell \Big\{ (m-\ell-K-1)(m-2\ell+2) \\
&- (m-\ell+1)(m-2\ell+1) \Big\} \Big\} \Big/ \prod_{k=m-\ell-(K+1)}^{m-\ell+1} k \\
=&(-1)^{K+2+\ell} \frac{K!(K+1)!}{\ell!(K+2-\ell)!} (K+2)(K+1)(m-2\ell+1) \Big/ \prod_{k=m-\ell-(K+1)}^{m-\ell+1} k
\end{aligned}$$

agreeing with (2.56). The proof of the induction step when $\ell = 0$ or $\ell = K+2$ is similar. Hence the result follows by induction. ∎

Lemma 2.31. *Let $n \in \mathbb{N}_0$ and $y \in \mathbb{R}^4$, $|y| < 1$. For $p \in \mathbb{N}$ let*

$$\widehat{g}_p(y) = \sum_{m=1}^{p} \sum_{\ell=0}^{\min\{\lfloor m/2 \rfloor, n+1\}} c^n_{m,\ell} |y|^{2\ell} \chi_{m-2\ell}(y). \tag{2.57}$$

If $p > 2n$ then

$$\Big| |I-y|^{2n} \ln |I-y|^2 - \widehat{g}_p(y) \Big| \le \frac{2^{n+1} n! (p+2)^2}{(p+1-n)\cdots(p-2n)} \frac{|y|^{p+1}}{1-|y|}.$$

Proof. By Lemma 2.29,

$$\Big| |I-y|^{2n} \ln |I-y|^2 - \widehat{g}_p(y) \Big| \le \left| \sum_{m=p+1}^{\infty} \sum_{\ell=0}^{n+1} c^n_{m,\ell} |y|^{2\ell} \chi_{m-2\ell}(y) \right|. \tag{2.58}$$

Then using Lemma 2.30 the magnitude of all the terms of order $|y|^m$ can be estimated by

$$\begin{aligned}\left|\sum_{\ell=0}^{n+1} c^n_{m,\ell}|y|^{2\ell}\chi_{m-2\ell}(y)\right| &\leq \sum_{\ell=0}^{n+1}\frac{(m-\ell-n-1)!}{(m-\ell+1)!}(m-2\ell+1)n!\binom{n+1}{\ell}(m-2\ell+1)|y|^m \\ &\leq n!|y|^m\sum_{\ell=0}^{n+1}\frac{(m-2\ell+1)^2}{(m-\ell+1)\cdots(m-\ell-n)}\binom{n+1}{\ell} \\ &\leq n!|y|^m q_n(m)\sum_{\ell=0}^{n+1}\binom{n+1}{\ell} \\ &= n!2^{n+1}|y|^m q_n(m),\end{aligned}$$

where

$$q_n(m) = \frac{(m+1)^2}{(m-n)\big(m-(n+1)\big)\cdots\big(m-(2n+1)\big)}. \tag{2.59}$$

The function $q_n(m)$ is positive and decreasing in m for $m > 2n+1$. Hence the right hand side of (2.58) can be estimated as

$$\begin{aligned}\left|\sum_{m=p+1}^{\infty}\sum_{\ell=0}^{n+1} c^n_{m,\ell}|y|^{2\ell}\chi_{m-2\ell}(y)\right| &\leq \sum_{m=p+1}^{\infty}\left|\sum_{\ell=0}^{n+1} c^n_{m,\ell}|y|^{2\ell}\chi_{m-2\ell}(y)\right| \\ &\leq \sum_{m=p+1}^{\infty} n!2^{n+1}|y|^m q_n(m) \\ &\leq 2^{n+1}n!q_n(p+1)\frac{|y|^{p+1}}{1-|y|}.\end{aligned}$$

∎

The full expansion and error bound

Combining the results in (2.49), Lemma 2.27 and Lemma 2.29, the function

$$\Phi_{x_<}(x) = |x-x_<|^{2n}\ln|x-x_<|^2$$

may be approximated by the truncated series

$$g_p(x) = \ln|x|^2\sum_{m=0}^{n}|x|^{2m}\sum_{\ell=0}^{n-m} b^n_{m,\ell}|x_<|^{2\ell}\chi_{n-m-\ell}(x^*x_<) + |x|^{2n}\widehat{g}_p(x^{-1}x_<),$$

where $\widehat{g}_p$ is defined in (2.57). Then from Lemma 2.31 we obtain the error bound

$$\left|\Phi_{x_<}(x) - g_p(x)\right| \leq |x|^{2n}\frac{2^{n+1}n!(p+2)^2}{(p+1-n)\cdots(p-2n)}\,\frac{|x^{-1}x_<|^{p+1}}{1-|x^{-1}x_<|},$$

for $|x| > |x_<|$. Since $p > 2n$ this bound is increasing in $|y| = |x_<|/|x|$. We can apply the bound to each centre x_k in turn and sum. This gives

Theorem 2.32. *Suppose $x_k \in \mathbb{R}^4$, $|x_k| \le r$ and $d_k \in \mathbb{R}$ for each $1 \le k \le N$. Let s be the $(n+2)$-harmonic spline*

$$s(x) = \sum_{k=1}^{N} d_k |x - x_k|^{2n} \ln |x - x_k|^2.$$

Furthermore, let $B_{m,\ell}$ be the $(n-m-\ell+1) \times (n-m-\ell+1)$ matrix

$$B_{m,\ell} = \left[B_{i,j}^{m,\ell} \right]_{i,j=0}^{n-m-\ell} = b_{m,\ell}^{n} \sum_{k=1}^{N} d_k |x_k|^{2\ell} \big(T_{n-m-\ell}(x_k^*) \big),$$

and $C_{m,\ell}$ be the $(m-2\ell+1) \times (m-2\ell+1)$ matrix

$$C_{m,\ell} = \left[C_{i,j}^{m,\ell} \right]_{i,j=0}^{n-m-\ell} = c_{m,\ell}^{n} \sum_{k=1}^{N} d_k |x_k|^{2\ell} \big(T_{m-2\ell}(x_k) \big),$$

where the coefficients $b_{m,\ell}^n$ and $c_{m,\ell}^n$ are given recursively by (2.51) and (2.55) respectively. Let $p \in \mathbb{N}$, $p > 2n$, and set

$$\begin{aligned} g_p(x) &= \ln |x|^2 \sum_{m=0}^{n} \sum_{\ell=0}^{n-m} \sum_{i,j=0}^{n-m-\ell} B_{j,i}^{m,\ell} |x|^{2m} t_{i,j}^{n-m-\ell}(x) \\ &\quad + \sum_{\ell=0}^{n+1} \sum_{m=\max\{1,2\ell\}}^{p} \sum_{i,j=0}^{m-2\ell} C_{j,i}^{m,\ell} |x|^{2(n+1-\ell)} o_{i,j}^{m-2\ell}(x) \\ &= \ln |x|^2 \sum_{m=0}^{n} |x|^{2m} \sum_{\ell=0}^{n-m} \operatorname{Tr} \left(B_{m,\ell} T_{n-m-\ell}(x) \right) \\ &\quad + \sum_{\ell=0}^{n+1} \sum_{m=\max\{1,2\ell\}}^{p} |x|^{2(n+1-\ell)} \operatorname{Tr} \big(C_{m,\ell} O_{m-2\ell}(x) \big), \end{aligned} \tag{2.60}$$

$x \in \mathbb{R}^4 \backslash \{0\}$. Then for all x with $|x| > r$

$$\big| s(x) - g_p(x) \big| \le M r^{2n} \frac{(p+2)^2 2^{n+1} n!}{(p+1-n) \cdots (p-2n)} \left(\frac{1}{c} \right)^{p-2n+1} \frac{1}{1 - 1/c},$$

where $M = \sum_{k=1}^{N} |d_k|$ and $c = |x|/r$.

2.4 Uniqueness

In this section we will prove that the truncated expansions, g_p, appearing in (2.46) and (2.60) are the only functions of these forms achieving the stated asymptotic accuracy in approximating s as $|x| \to \infty$. These uniqueness results will allow us to form far field expansions in an inexpensive indirect manner, knowing that the expansions so obtained are identical with, and enjoy the same error estimates as, the computationally expensive directly formed expansions.

Lemma 2.33. *Let $p \in \mathbb{N}_0$. Suppose a function $\widetilde{g}_p$ defined for $x \in \mathbb{R}^4\backslash\{0\}$ can be written in the form*

$$\begin{aligned}\widetilde{g}_p(x) = \ln|x|^2 \sum_{m=0}^{n}\sum_{\ell=0}^{n-m}\sum_{i,j=0}^{n-m-\ell} \widetilde{B}_{j,i}^{m,\ell}|x|^{2m} t_{i,j}^{n-m-\ell}(x) \\ + \sum_{\ell=0}^{n+1}\sum_{m=\max\{1,2\ell\}}^{p}\sum_{i,j=0}^{m-2\ell} \widetilde{C}_{j,i}^{m,\ell}|x|^{2(n+1-\ell)} o_{i,j}^{m-2\ell}(x),\end{aligned} \tag{2.61}$$

where the various coefficients are complex numbers. Then

(i) *The coefficients $\{\widetilde{B}_{j,i}^{m,\ell}\}$ and $\{\widetilde{C}_{j,i}^{m,\ell}\}$ are uniquely determined by the function $\widetilde{g}_p$.*

(ii) *If $p > 2n$ and*

$$|\widetilde{g}_p(x)| = o(|x|^{2n-p}), \qquad \text{as } |x| \to \infty,$$

then $\widetilde{g}_p$ is identically zero.

Proof. We will need to use the fundamental properties of the inner and outer functions developed in Section 2.2. Recall that $t_{i,j}^m$ is homogeneous of order m and $\{t_{i,j}^m : 0 \le i,j \le m,\ 0 \le m \le q\}$ is an orthogonal set of non-trivial spherical harmonics on the unit sphere $\mathcal{S}^3$. The definition of the outer functions (2.16)

$$o_{i,j}^m(z,w) = |x|^{-(2m+2)} t_{i,j}^m(\overline{z}, -w)$$

then implies that $o_{i,j}^m$ is homogeneous of order $-(m+2)$ and $\{o_{i,j}^m : 0 \le i,j \le m,\ 0 \le m \le q\}$ is linearly independent on $\mathcal{S}^3$.

Now fix $p \in \mathbb{N}_0$ and consider a function of the form (2.61). Rearrange the finite sum $\widetilde{g}_p$ by grouping together terms of the same growth at infinity, and arranging the groups in order of decreasing growth at infinity. The order of magnitude of $\widetilde{g}_p$ as $|x| \to \infty$ will be the same as that of the first non-zero group of terms.

Fix an integer k and denote the sum of the group of terms of growth $|x|^k \ln|x|$ by L_k. Thus

$$L_k(x) = \ln|x|^2 \sum_{m=0}^{n}\sum_{\ell=0}^{n-m}\sum_{i,j=0}^{n-m-\ell} \delta_{k,n+m-\ell}\widetilde{B}_{j,i}^{m,\ell}|x|^{2m} t_{i,j}^{n-m-\ell}(x).$$

Restricting attention to those terms for which the delta function is non-zero we see that among these a particular inner function $t_{i',j'}^{k'}$ can arise only when $k' = k - 2m$ and thus can arise at most once. Hence by the linear independence of $\{t_{i,j}^m : 0 \le i,j \le m,\ 0 \le m \le n\}$ on $\mathcal{S}^3$, $L_k(x)$ is identically zero for all $x \ne 0$ if and only if

$$\delta_{k,n+m-\ell}\widetilde{B}_{j,i}^{m,\ell} = 0,$$

for all $m = 0, \dots, n$; $\ell = 0, \dots, n-m$; $i, j = 0, \dots, n-m-\ell$. Similarly, the sum of the group of terms of growth $|x|^k$ at infinity is

$$G_k(x) = \sum_{\ell=0}^{n+1} \sum_{m=\max\{1,2\ell\}}^{p} \sum_{i,j=0}^{m-2\ell} \delta_{k,2n-m} \widetilde{C}_{j,i}^{m,\ell} |x|^{2(n+1-\ell)} o_{i,j}^{m-2\ell}.$$

Fix k and restrict attention to those terms for which the delta function is non zero. A particular outer function $o_{i',j'}^{m'}$ can arise only when $m' = 2n - k - 2\ell$. Since k and n are fixed this happens for at most one value of ℓ, and hence at most once. Thus by the linear independence of $\{o_{i',j'}^{m'} : 0 \le i', j' \le m', 0 \le m' \le q\}$ on $\mathcal{S}^3$, $G_k(x)$ is identically zero for all $x \ne 0$ if and only if

$$\delta_{k,2n-m} \widetilde{C}_{j,i}^{m,\ell} = 0,$$

for all $\ell = 0, \dots, n+1$; $m = \max\{1, 2\ell\}, \dots, p$; $i, j = 0, \dots, m$.

For (ii) just note that if $p > 2n$ then $o(|x|^{2n-p}) \to 0$ so no terms can appear in the $\ln |x|^2$ summand, as each of those terms do not decay. Each term in the second summand is homogeneous of order $2n - m \ge 2n - p$. But the decay rate $o(|x|^{2n-p})$ precludes these terms from occurring, *i.e.*, $\widetilde{g}_p = 0$. ∎

A simpler argument based on the same ideas shows

Lemma 2.34. *Let $p \in \mathbb{N}_0$. Suppose a function $\widetilde{g}_p$ defined for $x \in \mathbb{R}^4 \backslash \{0\}$ can be written in the form*

$$\widetilde{g}_p(x) = \sum_{m=0}^{p} \sum_{i,j=0}^{m} \widetilde{C}_{j,i}^{m} \, o_{i,j}^{m}(x).$$

Then

(i) *The coefficients $\{\widetilde{C}_{j,i}^{m}\}$ are uniquely determined by the function $\widetilde{g}_p$.*

(ii) *If*

$$|\widetilde{g}_p(x)| = o(|x|^{-(p+2)}), \qquad \textit{as } |x| \to \infty,$$

then $\widetilde{g}_p$ is identically zero.

2.5 Translation of expansions

In this section we develop formulae which enable us to obtain a truncated expansion about one centre indirectly from a truncated expansion about another. The operation count for this translation operation depends only on the order of the original expansion, not upon the number of centres x_k underlying it. In contrast, the operation count for direct expansion of a cluster, is $\mathcal{O}\big(N(n+1)p^3\big)$ where N is the number of centres in the cluster. Thus indirect formation of expansions can be more efficient than direct expansion when the number of centres in a particular cluster is large and truncated expansions of sub-clusters are available.

For any matrix $A = (a_{i',j'})$, denote by $A|_{i,j}^m$ the $(m+1)\times(m+1)$ sub-matrix of A which begins at the (i,j) position, *i.e.*,

$$(A_{i,j}^m)_{i',j'} = a_{i+i',j+j'}, \qquad i',j' = 0,\dots,m.$$

Theorem 2.35 (Outer to outer or inner translation). *Let $x, x_< \in \mathbb{H}_0^1$ be such that $0 < |x_<| < |x|$. Then*

$$o_{i,j}^m(x - x_<) = \sum_{m'=0}^{\infty} \operatorname{Tr}\left(E(m,m',i) O_{m+m'}(x)\Big|_{i,j}^{m'} T_{m'}(x_<)\right), \tag{2.62}$$

where $E(m,m',i)$ is the $(m'+1)\times(m'+1)$ diagonal matrix with entries

$$e(m,m',i)_{i',i'} = \binom{i+i'}{i}\binom{m+m'-(i+i')}{m-i}.$$

Proof. Using Lemma 2.23, Lemma 2.10 and the relationship (2.37) between the outer functions and the operators $R_{i,j}^m$,

$$\begin{aligned}
o_{i,j}^m(x - x_<) = R_{i,j}^m \frac{1}{|x-x_<|^2} &= R_{i,j}^m \sum_{m'=0}^{\infty} \operatorname{Tr}\Big(O_{m'}(x) T_{m'}(x_<)\Big) \\
&= R_{i,j}^m \sum_{m'=0}^{\infty} \sum_{i',j'=0}^{m'} o_{i',j'}^{m'}(x) t_{j',i'}^{m'}(x_<) \\
&= R_{i,j}^m \sum_{m'=0}^{\infty} \sum_{i',j'=0}^{m'} t_{j',i'}^{m'}(x_<) R_{i',j'}^{m'} \frac{1}{|x|^2} \\
&= \sum_{m'=0}^{\infty} \sum_{i',j'=0}^{m'} t_{j',i'}^{m'}(x_<) R_{i,j}^m R_{i',j'}^{m'} \frac{1}{|x|^2} \\
&= \sum_{m'=0}^{\infty} \sum_{i',j'=0}^{m'} t_{j',i'}^{m'}(x_<) e(m,m',i)_{i',i'} R_{i+i',j+j'}^{m+m'} \frac{1}{|x|^2} \\
&= \sum_{m'=0}^{\infty} \sum_{i',j'=0}^{m'} e(m,m',i)_{i',i'} o_{i+i',j+j'}^{m+m'}(x) t_{j',i'}^{m'}(x_<) \\
&= \sum_{m'=0}^{\infty} \operatorname{Tr}\left(E(m,m',i) O_{m+m'}(x)\Big|_{i,j}^{m'} T_{m'}(x_<)\right),
\end{aligned}$$

where the differentiation term by term is justified by the real analyticity. ∎

This Theorem is sufficient to translate a far field expansion of the type in (2.46). In particular consider an expansion like (2.46) but centred on $x_< \neq 0$. Then

$$\begin{aligned} g_p(x - x_<) &= \sum_{m=0}^{p} \sum_{i,j=0}^{m} C_{j,i}^m \, o_{i,j}^m(x - x_<) \\ &= \sum_{m=0}^{p} \sum_{i,j=0}^{m} C_{j,i}^m \sum_{m'=0}^{\infty} \sum_{i',j'=0}^{m'} e(m, m', i)_{i',i'} o_{i+i',j+j'}^{m+m'}(x) t_{j',i'}^{m'}(x_<) \\ &= \sum_{m=0}^{p} \sum_{i,j=0}^{m} D_{j,i}^m \, o_{i,j}^m(x) + \mathcal{O}\big(|x|^{-(p+3)}\big) \end{aligned} \tag{2.63}$$

where the coefficients $D_{j,i}^m$ are defined by the "convolution"

$$\binom{m}{i} D_{j,i}^m = \sum_{m'=0}^{m} \binom{m}{m'} \sum_{i',j'=0}^{m'} \binom{m-m'}{i-i'} t_{j-j',i-i'}^{m-m'}(x_<) \binom{m'}{i'} C_{j',i'}^{m'}. \tag{2.64}$$

Thus

$$h_p(x) = \sum_{m=0}^{p} \sum_{i,j=0}^{m} D_{j,i}^m \, o_{i,j}^m(x)$$

approximates $g_p(x)$ with error of order $\mathcal{O}(|x|^{-(p+3)})$ as $|x| \to \infty$. But by Theorem 2.24 the series formed directly, u_p, is of the same form and shares the same order of approximation. Hence the difference $u_p(x) - h_p(x)$ is $\mathcal{O}(|x|^{-(p+3)})$ as $|x| \to \infty$, and by the Uniqueness Theorem, Lemma 2.34, u_p and h_p are identical.

Furthermore, Theorem 2.35 is sufficient to translate an expansion of the form (2.63) into a Taylor series about 0 (a Maclaurin series). Since the functions $o_{i,j}^m$ are homogenous of degree m,

$$o_{i,j}^m(x - x_<) = (-1)^m o_{i,j}^m(x_< - x).$$

This simple observation allows the roles of x and $x_<$ to be switched in the application of Theorem 2.35. Starting with g_p defined by (2.63) and proceeding in this manner, we obtain

$$\begin{aligned} g_p(x - x_<) &= \sum_{m'=0}^{p} \sum_{i',j'=0}^{m'} C_{j',i'}^{m'} (-1)^{m'} \, o_{i',j'}^{m'}(x_< - x) \\ &= \sum_{m'=0}^{p} \sum_{i',j'=0}^{m'} (-1)^{m'} C_{j',i'}^{m'} \sum_{m=0}^{\infty} \sum_{i,j=0}^{m} e(m', m, i')_{i,i} o_{i+i',j+j'}^{m+m'}(x_<) t_{j,i}^m(x) \\ &= \sum_{m=0}^{p} \sum_{i,j=0}^{m} F_{j,i}^m \, t_{i,j}^m(x) + \mathcal{O}\big(|x|^{p+1}\big), \quad \text{as } |x| \to 0, \end{aligned}$$

where the coefficients $F^m_{j,i}$ are given by the "correlation"

$$\binom{m}{j}^{-1} F^m_{j,i} = \sum_{m'=0}^{p} \binom{m+m'}{m} \sum_{i',j'=0}^{m'} \binom{m+m'}{j+j'}^{-1} o^{m+m'}_{j+j',i+i'}(x_<)\,(-1)^{m'}\binom{m'}{j'} C^{m'}_{i',j'}. \quad (2.65)$$

Then by the characterisation of the Maclaurin polynomial q of degree p for a function f as the only polynomial of total degree p with

$$|f(x) - q(x)| = o|x|^p, \qquad \text{as } |x| \to 0,$$

it follows that

$$q(\cdot) = \sum_{m=0}^{p} \sum_{i,j=0}^{m} F^m_{j,i}\, t^m_{i,j}(\cdot). \quad (2.66)$$

Theorem 2.36 (Inner-to-inner translation formula). *For all $x, x_< \in \mathbb{R}^4$, $m \in \mathbb{N}_0$ and $0 \le i, j \le m$,*

$$t^m_{i,j}(x - x_<) = \sum_{m'=0}^{m} \sum_{\substack{i'=\max\{0,i-(m-m')\}\\ j'=\max\{0,j-(m-m')\}}}^{\substack{i'=\min\{i,m'\}\\ j'=\min\{j,m'\}}} (-1)^{m-m'} \binom{m-j}{m'-j'} \binom{j}{j'} t^{m-m'}_{i-i',j-j'}(x_<) t^{m'}_{i',j'}(x).$$

Proof. Because the functions $\{t^{m'}_{i',j'} : 0 \le i', j' \le m',\ 0 \le m' \le m\}$ form a basis for harmonic homogenous polynomials of degree at most m, and since $t^m_{i,j}(\cdot - x_<)$ is such a polynomial,

$$t^m_{i,j}(x - x_<) = \sum_{m'=0}^{m} \sum_{i',j'=0}^{m'} a^{m,m'}_{i,i',j,j'} t^{m'}_{i',j'}(x),$$

for some coefficients $a^{m,m'}_{i,i',j,j'}$ that depend on $x_<$. Applying the functionals $\lambda^{m''}_{i'',j''}$ of Lemma 2.22 to this expression gives

$$\binom{j}{j''} \binom{m-j}{m''-j''} t^{m-m''}_{i-i'',j-j''}(-x_<) = a^{m,m''}_{i,i'',j,j''}.$$

Since $(-1)^{m-m''}$ factors out by the homogeneity of $t^{m-m''}_{i-i'',j-j''}$, the result follows once we recall that the left hand side is zero unless $j'' \le j$, $j - j'' \le m - m''$, $i'' \le i$ and $i - i'' \le m - m''$. ∎

This theorem may be used to translate a polynomial expansion such as (2.66). For example, if

$$q(x) = \sum_{m'=0}^{p} \sum_{i',j'=0}^{m'} F^{m'}_{j',i'}\, t^{m'}_{i',j'}(x - x_<), \quad (2.67)$$

then by Theorem 2.36, we get

$$
\begin{aligned}
q(x) &= \sum_{m'=0}^{p} \sum_{i',j'=0}^{m'} F_{j',i'}^{m'} \sum_{m=0}^{m'} \sum_{i,j=0}^{m} (-1)^{m'-m} \binom{m'-j'}{m-j} \binom{j'}{j} t_{i'-i,j'-j}^{m'-m}(x_<) t_{i,j}^{m}(x) \\
&= \sum_{m=0}^{p} \sum_{i,j=0}^{m} G_{j,i}^{m}\, t_{i,j}^{m}(x),
\end{aligned}
$$

where the coefficients $G_{j,i}^{m}$ are given by the "convolution"

$$
\binom{m}{j}^{-1} G_{j,i}^{m} = \sum_{m'=0}^{p} \binom{m'}{m} \sum_{i',j'=0}^{m'} \binom{m'}{j'}^{-1} F_{j',i'}^{m'}\, (-1)^{m'-m} \binom{m'-m}{j'-j} t_{i'-i,j'-j}^{m'-m}(x_<). \tag{2.68}
$$

It should be noted that this is an exact recentering of the polynomial q.

Just as we were able to translate expansions of the form (2.63), we want to be able to translate expansions like (2.60). One of our tools will be formulae for the products of z, w, $\overline{z}$ or $\overline{w}$ with a single inner or single outer function. These multiplication rules are contained in Lemmas 2.37 and 2.38 below.

Lemma 2.37. *For $m \geq 0$ and $0 \leq i, j \leq m$,*

$$
\begin{aligned}
z\ o_{i,j}^{m}(z,w) &= \frac{i+1}{m+1}\ |x|^2 o_{i+1,j+1}^{m+1}(z,w) + \frac{m-j}{m+1}\ o_{i,j}^{m-1}(z,w), && \text{(2.69a)}\\
w\ o_{i,j}^{m}(z,w) &= -\frac{m+1-i}{m+1}\ |x|^2 o_{i,j+1}^{m+1}(z,w) + \frac{m-j}{m+1}\ o_{i-1,j}^{m-1}(z,w), && \text{(2.69b)}\\
\overline{z}\ o_{i,j}^{m}(z,w) &= \frac{m+1-i}{m+1}\ |x|^2 o_{i,j}^{m+1}(z,w) + \frac{j}{m+1}\ o_{i-1,j-1}^{m-1}(z,w), && \text{(2.69c)}\\
\overline{w}\ o_{i,j}^{m}(z,w) &= \frac{i+1}{m+1}\ |x|^2 o_{i+1,j}^{m+1}(z,w) - \frac{j}{m+1}\ o_{i,j-1}^{m-1}(z,w). && \text{(2.69d)}
\end{aligned}
$$

Proof. First assume $m > 0$. Differentiate (2.17) with respect to $\overline{z}$. For the right hand side we obtain

$$
\begin{aligned}
(m-j)z_1(z_1\overline{z}+z_2\overline{w})^{m-1-j}(z_1(-w)+z_2 z)^j &= (m-j)z_1|x|^{2m} \sum_{i=0}^{m-1} z_1^{m-1-i} z_2^{i} o_{i,j}^{m-1}(z,w) \\
&= (m-j)|x|^{2m} \sum_{i=0}^{m-1} z_1^{m-i} z_2^{i} o_{i,j}^{m-1}(z,w). \quad (2.70)
\end{aligned}
$$

Since $|x|^2 = z\overline{z} + w\overline{w}$, for the left hand side we have

$$
\begin{aligned}
&z(m+1)|x|^{2m} \sum_{i=0}^{m} z_1^{m-i} z_2^{i} o_{i,j}^{m}(z,w) + |x|^{2(m+1)} \sum_{i=0}^{m} z_1^{m-i} z_2^{i} \frac{\partial}{\partial \overline{z}} o_{i,j}^{m}(z,w) \\
&= z(m+1)|x|^{2m} \sum_{i=0}^{m} z_1^{m-i} z_2^{i} o_{i,j}^{m}(z,w) - |x|^{2(m+1)} \sum_{i=0}^{m} z_1^{m-i} z_2^{i} (i+1) o_{i+1,j+1}^{m+1}(z,w) \quad (2.71)
\end{aligned}
$$

where we have used (2.32d) to evaluate $\frac{\partial}{\partial \overline{z}} o^m_{i,j}$. By considering the coefficient of $z_1^{m-i} z_2^i$ in (2.70) and (2.71) we see that

$$(m-j) o^{m-1}_{i,j}(z,w) = z(m+1) o^m_{i,j}(z,w) - |x|^2 (i+1) o^{m+1}_{i+1,j+1}(z,w)$$

for $0 \leq i,j \leq m$, which proves (2.69a)

By differentiating (2.17) with respect to $\overline{w}$, z and w, in a similar manner we obtain (2.69b), (2.69c) and (2.69d) respectively, for $m > 0$. The special case of $m = 0$ for (2.69) follows directly from $o^0_{0,0} = |x|^{-2}$ and the recurrence relations (2.27). ∎

Substituting (2.16) into (2.69) leads to a similar result for the inner functions. Specifically,

Lemma 2.38. *For $m \geq 0$ and $0 \leq i,j \leq m$,*

$$z\ t^m_{i,j}(z,w) = \frac{m+1-i}{m+1}\ t^{m+1}_{i,j}(z,w) + \frac{j}{m+1}\ |x|^2 t^{m-1}_{i-1,j-1}(z,w), \tag{2.72a}$$

$$w\ t^m_{i,j}(z,w) = \frac{m+1-i}{m+1}\ t^{m+1}_{i,j+1}(z,w) - \frac{m-j}{m+1}\ |x|^2 t^{m-1}_{i-1,j}(z,w), \tag{2.72b}$$

$$\overline{z}\ t^m_{i,j}(z,w) = \frac{i+1}{m+1}\ t^{m+1}_{i+1,j+1}(z,w) + \frac{m-j}{m+1}\ |x|^2 t^{m-1}_{i,j}(z,w), \tag{2.72c}$$

$$\overline{w}\ t^m_{i,j}(z,w) = -\frac{i+1}{m+1}\ t^{m+1}_{i+1,j}(z,w) + \frac{j}{m+1}\ |x|^2 t^{m-1}_{i,j-1}(z,w). \tag{2.72d}$$

Since $t^m_{i,j} = 0$ and $o^m_{i,j} = 0$ if $m < 0$, multiple applications of Lemma 2.37 and Lemma 2.38 may be used to obtain

Corollary 2.39. *Let p be a given homogenous polynomial of degree m' in z, $\overline{z}$, w and $\overline{w}$. Then to each inner function $t^m_{i,j}$ there correspond constants $\{F^{\ell'}_{i',j'}\}$ and to each outer function $o^m_{i,j}$ there correspond constants $\{G^{\ell'}_{i',j'}\}$ such that*

$$p(z,\overline{z},w,\overline{w})\ t^m_{i,j}(x) = \sum_{\ell'=0}^{\min\{m',\lfloor (m+m')/2 \rfloor\}} |x|^{2\ell'} \sum_{i',j'=0}^{m+m'-2\ell} F^{\ell'}_{i',j'} t^{m+m'-2\ell'}_{i',j'}(x),$$

$$p(z,\overline{z},w,\overline{w})\ o^m_{i,j}(x) = \sum_{\ell'=0}^{\min\{m',\lfloor (m+m')/2 \rfloor\}} |x|^{2(m'-\ell')} \sum_{i',j'=0}^{m+m'-2\ell'} G^{\ell'}_{i',j'} o^{m+m'-2\ell'}_{i',j'}(x).$$

We now demonstrate how these results may be used to translate truncated a far field series, such as (2.60), due to a polyharmonic spline. Let g_p be such a series with centre of expansion at $x_<$, *i.e.*,

$$g_p(x) = \ln|x-x_<|^2 \sum_{\ell=0}^{n} |x-x_<|^{2\ell} \left\{ \sum_{m=0}^{n-\ell} \sum_{i,j=0}^{n-\ell-m} B^{\ell,m}_{j,i} t^{n-\ell-m}_{i,j}(x-x_<) \right\}$$
$$+ \sum_{\ell=0}^{n+1} |x-x_<|^{2(n+1-\ell)} \left\{ \sum_{m=\max\{1,2\ell\}}^{p} \sum_{i,j=0}^{m-2\ell} C^{\ell,m}_{j,i} o^{m-2\ell}_{i,j}(x-x_<) \right\}. \tag{2.73}$$

The translations of objects such as those in the two sets of curly braces has been discussed already. The terms in the first set may be translated via Theorem 2.36 in much the same way as (2.67) was translated. The terms in the second set of braces may be translated using Theorem 2.35 in a similar manner to (2.63). Let $\{\widetilde{B}_{j,i}^{\ell,m}\}$ and $\{\widetilde{C}_{j,i}^{\ell,m}\}$ be the translated coefficients. Then

$$\begin{aligned} g_p(x) &= \ln|x-x_<|^2 \sum_{\ell=0}^{n} |x-x_<|^{2\ell} \left\{ \sum_{m=0}^{n-\ell} \sum_{i,j=0}^{n-\ell-m} \widetilde{B}_{j,i}^{\ell,m} t_{i,j}^{n-\ell-m}(x) \right\} \\ &\quad + \sum_{\ell=0}^{n+1} |x-x_<|^{2(n+1-\ell)} \left\{ \sum_{m=\max\{1,2\ell\}}^{p} \sum_{i,j=0}^{m-2\ell} \widetilde{C}_{j,i}^{\ell,m} o_{i,j}^{m-2\ell}(x) \right\} + \mathcal{O}\big(|x|^{-(p+1-2n)}\big). \quad (2.74) \end{aligned}$$

Recall that

$$|x-x_<|^2 = |x|^2 - 2\langle x, x_<\rangle + |x_<|^2 = |x|^2 - (z\overline{z_<} + \overline{z}z_< + w\overline{w_<} + \overline{w}w_<) + |x_<|^2.$$

Thus we may use Lemma 2.38 to "translate" any product of the form $|x-x_<|^2 t_{i',j'}^{m'}(x)$ into a sum of at most ten terms of the form $|x|^{2\ell''} t_{i'',j''}^{m''}(x)$, where the coefficients of those terms depend on $x_<$. An analogous procedure employing Lemma 2.37 translates a product of the form $|x-x_<|^2 o_{i',j'}^{m'}(x)$ into a sum of at most ten terms of the form $|x|^{2\ell''} o_{i'',j''}^{m''}(x)$. Applying this procedure repeatedly to (2.74) viewed as a nested product

$$f_0(x) + |x-x_<|^2\Big(f_1(x) + |x-x_<|^2\Big(f_2(x) + \cdots + |x-x_<|^2\Big(f_n(x) + |x-x_<|^2 f_{n+1}(x)\Big)\cdots\Big)\Big)$$

brings g_p to the form

$$\begin{aligned} g_p(x) &= \ln|x-x_<|^2 \sum_{\ell=0}^{n} |x|^{2\ell} \sum_{m=0}^{n-\ell} \sum_{i,j=0}^{n-\ell-m} \approx{B}_{j,i}^{\ell,m} t_{i,j}^{n-\ell-m}(x) \\ &\quad + \sum_{\ell=0}^{n+1} |x|^{2(n+1-\ell)} \sum_{m=\max\{1,2\ell\}}^{p} \sum_{i,j=0}^{m-2\ell} \approx{C}_{j,i}^{\ell,m} o_{i,j}^{m-2\ell}(x) + \mathcal{O}\big(|x|^{-(p+1-2n)}\big), \quad (2.75) \end{aligned}$$

with only the $\ln|x-x_<|^2$ term left untranslated. This particular step in the translation costs $\mathcal{O}\big((n+1)^2(p+1)^3\big)$ operations. This is acceptable since n is small, typically $n \le 2$, and the number of terms in the series to be translated is $\mathcal{O}\big((n+1)(p+1)^3\big)$.

By Theorem 2.32,

$$\ln|x-x_<|^2 = \ln|x|^2 + \sum_{\ell'=0}^{1} |x|^{2-2\ell'} \sum_{m'=\max\{1,2\ell'\}}^{p'} \sum_{i',j'=0}^{m'-2\ell'} D_{j',i'}^{\ell',m'}(x_<) o_{i',j'}^{m'-2\ell'}(x) + \mathcal{O}\big(|x|^{-(p'+1)}\big).$$

Substituting this into (2.75) gives

$$g_p(x) = \ln|x|^2 \sum_{\ell=0}^{n} |x|^{2\ell} \sum_{m=0}^{n-\ell} \sum_{i,j=0}^{n-\ell-m} \widetilde{\widetilde{B}}_{j,i}^{\ell,m} \, t_{i,j}^{n-\ell-m}(x)$$
$$+ F_{x_<}(x) + \sum_{\ell=0}^{n+1} |x|^{2(n+1-\ell)} \sum_{m=\max\{1,2\ell\}}^{p} \sum_{i,j=0}^{m-2\ell} \widetilde{\widetilde{C}}_{j,i}^{\ell,m} \, o_{i,j}^{m-2\ell}(x) + \mathcal{O}\big(|x|^{-(p+1-2n)}\big), \quad (2.76)$$

where $F_{x_<}(x)$ is given by the product

$$\left\{ \sum_{\ell=0}^{n} |x|^{2\ell} \sum_{m=0}^{n-\ell} \sum_{i,j=0}^{n-\ell-m} \widetilde{\widetilde{B}}_{j,i}^{\ell,m} \, t_{i,j}^{n-\ell-m}(x) \right\}$$
$$\times \left\{ \sum_{\ell'=0}^{1} |x|^{2-2\ell'} \sum_{m'=\max\{1,2\ell'\}}^{p'} \sum_{i',j'=0}^{m'-2\ell'} D_{j',i'}^{\ell',m'}(x_<) o_{i',j'}^{m'-2\ell'}(x) \right\} \quad (2.77)$$

after it has been truncated by removing terms that are $\mathcal{O}(|x|^{-(p+1-2n)})$ as $|x| \to \infty$. From Corollary 2.39, (2.77) can be written in the form

$$\sum_{\ell=0}^{n+1} |x|^{2(n+1-\ell)} \sum_{m=\max\{1,2\ell\}}^{p} \sum_{i,j=0}^{m-2\ell} \widetilde{D}_{j,i}^{\ell,m} o_{i,j}^{m-2\ell}(x).$$

This completes the translation.

Chapter 3

Generalised multiquadrics in $\mathbb{R}^n$

Multiquadrics are a popular choice of radial basis function for interpolating scattered data in one or more dimensions. Many applications are described in the literature including geodesy, image processing and natural resource modelling (see, for example, Hardy [33]).

This chapter provides appropriate results for the application of the Fast Multipole Method to generalised multiquadric radial basis functions in $\mathbb{R}^n$. That is for functions of the form

$$s(x) = \sum_{i=1}^{N} d_i \Phi(x - t_i; k, \tau), \tag{3.1}$$

where

$$\Phi(x) = \Phi(x; k, \tau) = \left(x^2 + \tau^2\right)^{k/2}, \tag{3.2}$$

k is an odd integer, $\tau \geq 0$ and $x \in \mathbb{R}^n$. Note we will usually use the notation $\Phi(x)$ which hides the dependence of Φ on k and τ. The derived series and the analysis also apply when τ varies, that is, when the multiquadric parameter τ changes with the centre t_i.

This chapter is laid out as follows. First Sections 3.1 and 3.2 derive far field expansions of the following form

$$\Phi(x - t; k, \tau) = \sum_{\ell=0}^{\infty} P_\ell^{(k)}(|t|^2 + \tau^2, -2\langle t, x\rangle, |x|^2)/|x|^{2\ell-k} \tag{3.3}$$

where the $P_\ell^{(k)}$ are the polynomials

$$P_\ell(a,b,c) = P_\ell^{(k)}(a,b,c) = \sum_{j=\lfloor\frac{\ell+1}{2}\rfloor}^{\ell} \binom{k/2}{j}\binom{j}{\ell-j} b^{2j-\ell}(ac)^{\ell-j}, \qquad \ell \geq 0, \tag{3.4}$$

and $P_\ell^{(k)}$ is the zero function for negative ℓ. Section 3.2 also gives error bounds on approximations formed by truncating the series. Section 3.3 proves the uniqueness of the expansions that allows for parental expansions to be calculated by translating the child expansions as described in Section 1.2.2. Section 3.4 discusses recurrence relations for the efficient direct calculation of the far field coefficients. It shows that the terms of the first $p+k+1$ homogeneous orders in the series for an m centre cluster can be calculated in $\mathcal{O}(mn(p+k)^n)$ flops. Section 3.5 sets up some machinery which is used in Section 3.6 to establish methods for indirectly translating far field expansions. Section 3.7 shows how to efficiently convert a far field expansion into a local polynomial approximation. The chapter concludes with some numerical results showing that multiquadric radial basis functions can indeed be evaluated using this approach at a cost that grows as $\mathcal{O}(N \log N)$ in the number N of centres.

We will use lower case ϕ for the basic function as a function of one variable and upper case Φ for the function of n variables, *i.e.*, $\Phi = \phi(|\cdot|)$. It is common for the constant in the multiquadric basic function to be represented by c. However, we will use τ for this purpose, *i.e.*, the ordinary multiquadric basic function will be $\phi(r) = \sqrt{r^2+\tau^2}$. In the far field expansions, x is the *evaluation point* and u is the *centre of expansion*, although often we may take $u=0$. This centre of expansion should not be confused with the *centres* $\{t_i\}$ which are the centres of the radially symmetric components in the RBF. In applications these centres will often be the nodes of interpolation or point sources of some potential field.

3.1 A generating function

In this section we develop some important properties of the functions

$$f_k(z) = (\sqrt{az^2+bz+c})^k, \qquad k \in \mathbb{Z} \text{ is odd.} \tag{3.5}$$

These functions will turn out to be the generating functions for the polynomials $P_\ell^{(k)}$ that occur in the far and near field expansions of the generalised multiquadric function.

To fully explore the expansions of f_k we will need to use Gauss's hypergeometric function.

Lemma 3.1. *The* hypergeometric function*,*

$$F(a,b;c;z) = F(b,a;c;z) := \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{n=0}^{\infty} \frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)} \frac{z^n}{n!},$$

satisfies

$$F(a,b;c;z) = (1-z)^{c-a-b} F(c-a,c-b;c;z), \tag{3.6a}$$

$$\frac{d}{dz}F(a,b;c;z) = \frac{ab}{c} F(a+1,b+1;c+1;z). \tag{3.6b}$$

Furthermore, if a or b is equal to $-m$, m a non-negative integer, then $F(a, b; c; z)$ reduces to a polynomial of degree m in z.

Proof. See [1, Ch. 15]. ∎

Lemma 3.2. *Let $m, p \in \mathbb{N}_0$ and $|h| < 1$. Then*

$$\sum_{n=p}^{\infty} \binom{n+m}{m} h^n = \frac{h^p}{(1-h)^{m+1}} \frac{(p+m)!}{p!\, m!} F(-m, p; p+1; h).$$

Proof.

$$\begin{aligned}\sum_{n=p}^{\infty} \binom{n+m}{m} h^n &= \frac{(p+m)!}{p!\, m!} \frac{p!}{(p+m)!} h^p \sum_{n=0}^{\infty} \frac{(n+p+m)! n!}{(n+p)!} \frac{h^n}{n!} \\ &= \frac{(p+m)!}{p!\, m!} h^p F(m+p+1, 1; p+1; h) \\ &= \frac{(p+m)!}{p!\, m!} h^p (1-h)^{-(m+1)} F(-m, p; p+1; h),\end{aligned}$$

where the last equality follows from (3.6a). ∎

Lemma 3.3. (Churchill & Brown [21, pp. 127–128]) *Let $C = \{w \in \mathbb{C} : |w| = \rho\}$. If f is analytic inside and on C then for $|z| < \rho$,*

$$\left|f(z) - \big(T_\nu f\big)(z)\right| \leq \max_{w \in C} |f(w)| \left(\frac{|z|}{\rho}\right)^{\nu+1} \frac{1}{1 - |z|/\rho},$$

where $T_\nu f$ is the Maclaurin polynomial of f of degree ν.

We now present the major result of this section which gives a series expansion for f_k and a bound for the error in approximating f_k by a truncation of this series.

Lemma 3.4. *Let $k \in \mathbb{Z}$ be odd and let $a, b, c \in \mathbb{R}$, with $a, c > 0$ and $b^2 \leq 4ac$. Then for all $z \in \mathbb{C}$ such that $|z| < \sqrt{c/a}$,*

$$f_k(z) = (\sqrt{az^2 + bz + c})^k = c^{k/2} \sum_{\ell=0}^{\infty} \left(\frac{z}{c}\right)^{\ell} P_\ell^{(k)}(a, b, c) \tag{3.7}$$

where the $P_\ell^{(k)}$ are the polynomials defined in Equation (3.4). Moreover, for all z such that $|z| < \sqrt{c/a}$ and $\nu \in \mathbb{N}$,

$$\left|(az^2+bz+c)^{k/2} - c^{k/2}\sum_{\ell=0}^{\nu}\left(\frac{z}{c}\right)^{\ell} P_\ell^{(k)}(a,b,c)\right|$$
$$\leq \begin{cases} 2^k c^{k/2}\left(\dfrac{|z|}{\sqrt{c/a}}\right)^{\nu+1}\dfrac{\sqrt{c/a}}{\sqrt{c/a}-|z|}, & \text{if } k>0 \\ \dbinom{\nu-k}{\nu+1} c^{k/2}\left(\dfrac{|z|}{\sqrt{c/a}}\right)^{\nu+1}\left(\dfrac{\sqrt{c/a}}{\sqrt{c/a}-|z|}\right)^{-k} & \\ \qquad\times F\left(k+1,\nu+1;\nu+2;\dfrac{z}{\sqrt{c/a}}\right), & \text{if } k<0. \end{cases}$$

Proof. Let $\sqrt{\cdot}$ denote the principal branch of the complex square root. Then f_k is analytic whenever $q(z) = az^2 + bz + c$ is away from the branch cut, *i.e.*, whenever $q(z)$ is not a non-positive real. Completing the square,

$$q(z) = a\left\{\left(z+\frac{b}{2a}\right)^2 + \frac{4ac-b^2}{4a^2}\right\},$$

and since $b^2 \leq 4ac$, it is easily seen that f_k is analytic away from

$$\left\{z = -\frac{b}{2a} + \mathrm{i}y : y \in \mathbb{R} \text{ and } |y| \geq \sqrt{\frac{4ac-b^2}{4a^2}}\right\}.$$

Hence f_k is analytic on the disc

$$D = D_\epsilon = \left\{z \in \mathbb{C} : |z| \leq \rho = (1-\epsilon)\sqrt{c/a}\right\}, \qquad 0<\epsilon<1.$$

For all sufficiently small $|z|$, two applications of the Binomial Theorem and some reordering gives

$$\begin{aligned} f_k(z) &= c^{k/2}\left(1+\frac{bz+az^2}{c}\right)^{k/2} \\ &= c^{k/2}\sum_{j=0}^{\infty}\binom{k/2}{j}\left(\frac{bz+az^2}{c}\right)^j \\ &= c^{k/2}\sum_{j=0}^{\infty}\binom{k/2}{j}\sum_{q=0}^{j}\binom{j}{q}\frac{(bz)^{j-q}(az^2)^q}{c^j} \\ &= c^{k/2}\sum_{\ell=0}^{\infty}\sum_{j=\lfloor\frac{\ell+1}{2}\rfloor}^{\ell}\binom{k/2}{j}\binom{j}{\ell-j}\frac{(bz)^{2j-\ell}(az^2)^{\ell-j}}{c^j} \\ &= c^{k/2}\sum_{\ell=0}^{\infty}\left(\frac{z}{c}\right)^{\ell}\sum_{j=\lfloor\frac{\ell+1}{2}\rfloor}^{\ell}\binom{k/2}{j}\binom{j}{\ell-j}b^{2j-\ell}(ac)^{\ell-j} \end{aligned}$$

$$= c^{k/2}\sum_{\ell=0}^{\infty}\left(\frac{z}{c}\right)^{\ell} P_{\ell}^{(k)}(a,b,c).$$

Since the reordering of the double sum is valid for $|z|$ suffciently small, for such z this is the MacLaurin series for f_k. This now relation extends to all of D by the uniqueness of the MacLaurin series of f_k, proving the first part of the Lemma.

We will prove the second part separately for $k > 0$ and $k < 0$. For $k > 0$ we will apply the well known bound for the error in Taylor polynomial approximation given in Lemma 3.3. Fix z with $|z| < \sqrt{c/a}$ and choose ϵ with $0 < \epsilon < 1$ so small that $z \in D_\epsilon$. We apply the bound with $C = \partial D_\epsilon$. Firstly, note that

$$q(z) = a(z-\xi_+)(z-\xi_-), \qquad \xi_{\pm} = \frac{-b \pm \sqrt{b^2-4ac}}{2a},$$

and that both roots of q are outside D. Since $f_k(z) = q(z)^{k/2}$,

$$\max_{w\in C}|f_k(w)| = \left(\max_{w\in C}|q(w)|\right)^{k/2}.$$

For $w \in \partial D$,

$$|w-\xi_\pm| \le |w| + |\xi_\pm| = \rho + \sqrt{c/a} < 2\sqrt{c/a},$$

and thus

$$\max_{w\in\partial D}|q(w)| = |a|\max_{w\in\partial D}\{|w-\xi_+||w-\xi_-|\} \le |a|\left(2\sqrt{c/a}\right)^2 = 4c.$$

Now applying Lemma 3.3,

$$\begin{aligned}
&\left|(az^2+bz+c)^{k/2} - c^{k/2}\sum_{\ell=0}^{\nu}\left(\frac{z}{c}\right)^{\ell} P_{\ell}^{(k)}(a,b,c)\right| \\
&\qquad \le \max_{w\in\partial D}|f_k(w)|\left(\frac{|z|}{\rho}\right)^{\nu+1}\frac{1}{1-|z|/\rho} \\
&\qquad \le (4c)^{k/2}\left(\frac{|z|}{(1-\epsilon)\sqrt{c/a}}\right)^{\nu+1}\frac{(1-\epsilon)\sqrt{c/a}}{(1-\epsilon)\sqrt{c/a}-|z|}.
\end{aligned}$$

Taking the limit as ϵ goes to zero from above gives the result for $k > 0$.

For the case $k < 0$, write the polynomial q in the form

$$\begin{aligned}
q(z) =& az^2+bz+c = c\left\{1 + \frac{b}{\sqrt{ac}}\left(\frac{z}{\sqrt{c/a}}\right) + \left(\frac{z}{\sqrt{c/a}}\right)^2\right\} \\
=& c\left(1-2x\xi+\xi^2\right),
\end{aligned}$$

where

$$x = -\frac{1}{2}\frac{b}{\sqrt{ac}} \qquad \text{and} \qquad \xi = \frac{z}{\sqrt{c/a}}.$$

Now recall [64, (4.7.23)] that $\left(1 - 2x\xi + \xi^2\right)^{-\lambda}$ is the generating function for the *Gegenbauer (or ultraspherical) polynomials* $C_\ell^{(\lambda)}(x)$, *i.e.*,

$$\sum_{\ell=0}^{\infty} C_\ell^{(\lambda)}(x)\xi^\ell = (1 - 2x\xi + \xi^2)^{-\lambda}.$$

Letting $\lambda = -k/2$, we see that

$$f_k(z) = c^{k/2} \sum_{\ell=0}^{\infty} C_\ell^{(\lambda)}(x)\xi^\ell,$$

and thus equating coefficients

$$\left(\frac{z}{c}\right)^\ell P_\ell^{(k)}(a, b, c) = C_\ell^{(\lambda)}(x)\xi^\ell, \qquad \ell \in \mathbb{N}_0. \tag{3.8}$$

For $-1 \le x \le 1$,

$$\left|C_n^{(\lambda)}(x)\right| \le \binom{n + 2\lambda - 1}{n}, \qquad \lambda > 0,$$

[1, 22.14.2]. By the statement of the lemma, $b^2 \le 4ac$ and $|z| < \sqrt{c/a}$. This means that $-1 \le x \le 1$ and $|\xi| < 1$ and thus

$$\begin{aligned}\left|f_k(z) - c^{k/2} \sum_{\ell=0}^{\nu} \left(\frac{z}{c}\right)^\ell P_\ell^{(k)}(a, b, c)\right| &= \left|f_k(z) - c^{k/2} \sum_{\ell=0}^{\nu} C_\ell^{(-k/2)}(x)\xi^\ell\right| \\ &\le c^{k/2} \sum_{\ell=\nu+1}^{\infty} \binom{\ell - k - 1}{\ell} |\xi|^\ell \end{aligned} \tag{3.9}$$

By Lemma 3.2,

$$\sum_{\ell=\nu+1}^{\infty} \binom{\ell - k - 1}{\ell} |\xi|^\ell = \binom{\nu - k}{\nu + 1} \frac{|\xi|^{\nu+1}}{(1 - |\xi|)^{-k}} F\left(k + 1, \nu + 1; \nu + 2; |\xi|\right).$$

Using this in (3.9) we have

$$\begin{aligned}&\left|f_k(z) - c^{k/2} \sum_{\ell=0}^{\nu} \left(\frac{z}{c}\right)^\ell P_\ell^{(k)}(a, b, c)\right| \\ &\qquad \le c^{k/2} \binom{\nu - k}{\nu + 1} \left(\frac{|z|}{\sqrt{c/a}}\right)^{\nu+1} \left(\frac{\sqrt{c/a}}{\sqrt{c/a} - |z|}\right)^{-k} \\ &\qquad\qquad \times F\left(k + 1, \nu + 1; \nu + 2; \frac{z}{\sqrt{c/a}}\right).\end{aligned}$$

∎

In the case $k=-1$, the polynomial $F(k+1,\nu+1;\nu+2;z/\sqrt{c/a})$ that appears in the error bound of Lemma 3.4 is constant and has value 1. For all other negative values of k consider the function $F(k+1,p+1;p+2;\cdot)$ where $p\in\mathbb{N}_0$. Rephrasing Lemma 3.2 as

$$F(k+1,p;p+1;z)=\frac{p!\,(-k-1)!}{(p-k-1)!}\frac{(1-z)^{-k}}{z^p}\sum_{n=p}^{\infty}\binom{n-k-1}{-k-1}z^n,$$

it is easily seen that $F(k+1,p+1;p+2;\cdot)$ is non-negative on $[0,1)$. Using (3.6b) to differentiate F, we see that for $z\in[0,1)$

$$\frac{d}{dz}F(k+1,p;p+1;z)=\frac{(k+1)p}{p+1}F(k+2,p+1;p+2;z)\leq 0,$$

since $k<-1$. Since $F(\cdot,\cdot;\cdot;0)=1$, it follows that

$$F(k+1,\nu+1;\nu+2;z/\sqrt{c/a})\leq 1,\qquad k\in\mathbb{Z}_-,\quad |z|\leq\sqrt{c/a}. \tag{3.10}$$

As was observed in the proof of Lemma 3.4 and particularly in Equation (3.8), for $k<0$ the polynomials $P_\ell^{(k)}$ are closely related to the Gegenbauer polynomials $C_\ell^{(\lambda)}$ with $\lambda=-k/2$. However, many properties of the Gegenbauer polynomials are derived using their orthogonality with respect to the weight function $w(x)=(1-x^2)^{\lambda-1/2}$. This function is not integrable over the interval $[-1,1]$ when $\lambda\leq-1/2$, and thus we are unable to exploit properties of the Gegenbauer polynomials derived from orthogonality when $k\geq 1$. The following lemma can be identified as a well known recurrence for the Gegenbauer polynomials with parameter $\lambda=-k/2>-1/2$. Our proof here is based on the characterisation (3.7) and hence holds for all odd integers k.

Lemma 3.5. *Let $k\in\mathbb{Z}$ be odd. Then the polynomials $P_\ell^{(k)}$ defined in (3.4), satisfy the following recurrence relation for all $a,b,c\in\mathbb{R}$, and $\ell\in\mathbb{N}$:*

$$(\ell+1)P_{\ell+1}^{(k)}(a,b,c)=\left(\frac{k}{2}-\ell\right)bP_\ell^{(k)}(a,b,c)+\big(k-(\ell-1)\big)acP_{\ell-1}^{(k)}(a,b,c). \tag{3.11}$$

Proof. We will first prove the identity under the additional assumptions $a,c>0$, and $b^2\leq 4ac$. Making these assumptions and differentiating the right hand side of (3.7) term by term gives

$$f_k'(z)=c^{(k-2)/2}\sum_{\ell=0}^{\infty}\left(\frac{z}{c}\right)^\ell(\ell+1)P_{\ell+1}^{(k)}(a,b,c), \tag{3.12}$$

the term by term differentiation being valid for $|z|<\sqrt{c/a}$.

On the other hand differentiating the expression $f_k(z)=(\sqrt{az^2+bz+c})^k$ then expanding gives

$$f_k'(z)=\frac{k}{2}(az^2+bz+c)^{(k-2)/2}(2az+b)$$

$$
\begin{aligned}
&= \frac{k}{2} f_{k-2}(z)(2az+b) \\
&= \frac{k}{2} c^{(k-2)/2} \sum_{\ell=0}^{\infty} \left(\frac{z}{c}\right)^{\ell} (2az+b) P_{\ell}^{(k-2)}(a,b,c) \\
&= c^{(k-2)/2} \left\{ \sum_{\ell=0}^{\infty} \left(\frac{z}{c}\right)^{\ell} \frac{k}{2} b P_{\ell}^{(k-2)}(a,b,c) + \sum_{\ell=1}^{\infty} \left(\frac{z}{c}\right)^{\ell} kac P_{\ell-1}^{(k-2)}(a,b,c) \right\}. \qquad (3.13)
\end{aligned}
$$

Equating (3.12) and (3.13), then comparing coefficients gives

$$
(\ell+1) P_{\ell+1}^{(k)}(a,b,c) = \frac{k}{2} b P_{\ell}^{(k-2)}(a,b,c) + kac P_{\ell-1}^{(k-2)}(a,b,c), \quad \ell \in \mathbb{N}. \qquad (3.14)
$$

Using the obvious recurrence on f_k and then expanding gives

$$
\begin{aligned}
f_k(z) &= (az^2+bz+c) f_{k-2}(z) \\
&= c^{k/2} \left\{ \sum_{\ell=2}^{\infty} \left(\frac{z}{c}\right)^{\ell} ac P_{\ell-2}^{(k-2)}(a,b,c) + \sum_{\ell=1}^{\infty} \left(\frac{z}{c}\right)^{\ell} b P_{\ell-1}^{(k-2)}(a,b,c) \right. \\
&\qquad \left. + \sum_{\ell=0}^{\infty} \left(\frac{z}{c}\right)^{\ell} P_{\ell}^{(k-2)}(a,b,c) \right\}. \qquad (3.15)
\end{aligned}
$$

Equating (3.7) and (3.15), then comparing coefficients gives

$$
P_{\ell+1}^{(k)}(a,b,c) = P_{\ell+1}^{(k-2)}(a,b,c) + b P_{\ell}^{(k-2)}(a,b,c) + ac P_{\ell-1}^{(k-2)}(a,b,c), \quad \ell \in \mathbb{N}. \qquad (3.16)
$$

To obtain (3.11), multiply (3.16) by $(\ell+1)$ and equate to (3.14). Solving for $P_{\ell+1}^{(k-2)}(a,b,c)$ and making the index change $(k-2) \mapsto k$ gives (3.11).

This completes the proof when $a, c > 0$ and $b^2 \le 4ac$. This set in $\mathbb{R}^3$ contains a non-trivial open ball and polynomials in n variables are determined everywhere by their behaviour on any non-trivial open ball in $\mathbb{R}^n$. Hence (3.11) holds for all $a, b, c \in \mathbb{R}$ since the right and left hand sides of (3.11) are polynomial. ∎

3.2 Multivariate expansions.

Let $\Phi(x) = \left(x^2+\tau^2\right)^{k/2}$, where $\tau \ge 0$ and $k \in \mathbb{Z}$ is odd and where we have used the notational convenience $x^2 = \langle x, x\rangle = |x|^2$ for $x \in \mathbb{R}^n$. The following result gives a far field expansion for $\Phi(x-t)$ considered as a function of x, together with an error estimate for approximation with truncations of this expansion. The numerator polynomials $P_{\ell}^{(k)}(t^2+\tau^2, -2\langle t, x\rangle, x^2)$ that feature in the expansion are homogeneous of degree ℓ in x. Correspondingly, the ℓth term in the expansion is homogeneous of degree $k-\ell$ in x.

Lemma 3.6. *Let $k \in \mathbb{Z}$ be odd, $t \in \mathbb{R}^n$ and $\tau \geq 0$. For all $x \in \mathbb{R}^n$ with $|x| > \sqrt{t^2+\tau^2}$,*

$$\Phi(x-t) = \left((x-t)^2+\tau^2\right)^{k/2} = \sum_{\ell=0}^{\infty} P_\ell^{(k)}(t^2+\tau^2, -2\langle t, x\rangle, x^2)/|x|^{2\ell-k}$$

where the polynomials $P_\ell^{(k)}$ are defined in Equation (3.4). Moreover, for all x such that $|x| > \sqrt{t^2+\tau^2}$, and for all $p \in \mathbb{N}$ such that $p+k > 0$,

$$\left|\Phi(x-t) - \sum_{\ell=0}^{p+k} P_\ell^{(k)}(t^2+\tau^2, -2\langle t, x\rangle, x^2)/|x|^{2\ell-k}\right|$$
$$\leq \begin{cases} \left(2\sqrt{t^2+\tau^2}\right)^k \left(\dfrac{\sqrt{t^2+\tau^2}}{|x|}\right)^{p+1} \dfrac{|x|}{|x|-\sqrt{t^2+\tau^2}}, & \text{if } k > 0 \\ \dbinom{p}{p+k+1} \left(\sqrt{t^2+\tau^2}\right)^k \left(\dfrac{\sqrt{t^2+\tau^2}}{|x|}\right)^{p+1} & \\ \qquad \times \left(\dfrac{|x|}{|x|-\sqrt{t^2+\tau^2}}\right)^{-k}, & \text{if } k < 0. \end{cases}$$

Proof. Consider firstly the case when $\tau > 0$. Let $a = t^2+\tau^2$, $b = -2\langle t, x\rangle$ and $c = x^2$. Then

$$\Phi(x-t) = \left(x^2 - 2\langle t, x\rangle + t^2 + \tau^2\right)^{k/2} = f_k(1),$$

where f_k is the function defined in (3.5). Since $a, c > 0$, $b^2 \leq 4ac$, and $1 = |z| < \sqrt{c/a} = |x|/\sqrt{t^2+\tau^2}$, Lemma 3.4 may be applied with $\nu = p+k$ to yield the desired results when we recall the bound on F given by Equation (3.10).

This completes the proof when $\tau > 0$. For the remaining case fix x with $|x| > |t|$. Note that $0 < \widetilde{\tau} < \sqrt{|x|^2 - |t|^2}$ implies $|x| > \sqrt{t^2+\widetilde{\tau}^2}$. Hence the previous case can be applied to the expansion of

$$\Phi(x-t; k, \widetilde{\tau}) = \left((x-t)^2 + \widetilde{\tau}^2\right)^{k/2}$$

for all sufficiently small positive $\widetilde{\tau}$. Taking the limit as $\widetilde{\tau}$ goes to zero from above, and using the continuity of all the relevant quantities as functions of $\widetilde{\tau}$, gives the result for $\tau = 0$. ∎

Example 3.7. In the 1-dimensional case it is convenient to rewrite the series in the simpler form

$$\Phi(x-t) = \operatorname{sign}(x) \sum_{\ell=0}^{\infty} P_\ell^{(k)}(t^2+\tau^2, -2t, 1)/x^{\ell-k},$$

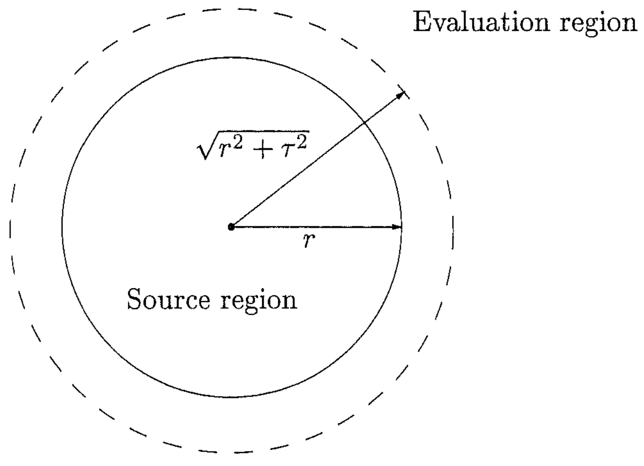


Figure 3.1: Region of validity of the far field expansion of a cluster.

which becomes, in the important special case ($k = 1$) of the ordinary multiquadric,

$$\sqrt{(x-t)^2+\tau^2} = \operatorname{sign}(x)\Big\{ x - t + \frac{1}{2}\tau^2 x^{-1} + \frac{1}{2}t\tau^2 x^{-2} + \frac{1}{8}(4t^2\tau^2 - \tau^4)x^{-3} + \frac{1}{8}(4t^3\tau^2 - 3t\tau^4)x^{-4} + \frac{1}{16}\tau^2(8t^4 - 12t^2\tau^2 + \tau^4)x^{-5} + \cdots + q_\ell(t,\tau)x^{1-\ell} + \cdots \Big\}.$$

Example 3.8. To display the componentwise form of the expansion in two dimensions we will temporarily adopt the notation $x = (x_1, x_2)$ and $t = (t_1, t_2)$. The far field expansion about zero of a single ordinary multiquadric basic function centred at t is then

$$\sqrt{|x-t|^2+\tau^2} = |x| - \frac{t_1x_1 + t_2x_2}{|x|} + \frac{1}{2}\frac{(t_2^2+\tau^2)x_1^2 + (t_1^2+\tau^2)x_2^2 - 2t_1t_2x_1x_2}{|x|^3} + \frac{1}{2}\frac{(t_1x_1+t_2x_2)\left\{(t_2^2+\tau^2)x_1^2 + (t_1^2+\tau^2)x_2^2 - 2t_1t_2x_1x_2\right\}}{|x|^5} + \cdots.$$

Since the bound of Lemma 3.6 is increasing in $|t|$ we can apply it to each centre in a cluster and sum obtaining the following expansion of the generalised multiquadric radial basis function associated with a cluster of centres. The geometry of the source cluster and the evaluation region is shown in Figure 3.1 above.

Theorem 3.9. *Suppose $t_i \in \mathbb{R}^n$, $|t_i| \le r$ and $d_i \in \mathbb{R}$ for each $1 \le i \le N$. Let k be odd, $\tau \ge 0$, and s be the generalised multiquadric spline*

$$s(x) = \sum_{i=1}^{N} d_j\Phi(x-t_i) = \sum_{i=1}^{N} d_i\left(\sqrt{(x-t_i)^2+\tau^2}\right)^k.$$

If $P_\ell^{(k)}$, $\ell \in \mathbb{N}_0$, are the polynomials defined by Equation (3.4), then the polynomials

$$Q_\ell(x) = \sum_{i=1}^{N} d_i P_\ell^{(k)}(t_i^2 + \tau^2, -2\langle t_i, x\rangle, x^2) \qquad \ell \in \mathbb{N}_0,$$

have the following property: Let $p \in \mathbb{N}_0$ and set

$$s_p(x) = \sum_{\ell=0}^{p+k} Q_\ell(x)/|x|^{2\ell-k}, \tag{3.17}$$

$x \in \mathbb{R}^n \backslash \{0\}$. Then for all x with $|x| > R = \sqrt{r^2+\tau^2}$

$$\left|s(x) - s_p(x)\right| \leq \begin{cases} 2^k M R^k \left(\frac{1}{c}\right)^{p+1} \frac{1}{1-1/c}, & \text{if } k > 0 \\ \binom{p}{p+k+1} M R^k \left(\frac{1}{c}\right)^{p+1} \left(\frac{1}{1-1/c}\right)^{-k}, & \text{if } k < 0, \end{cases}$$

where $M = \sum_{i=1}^{N} |d_i|$ and $c = |x|/R$.

3.3 The uniqueness of expansions

The uniqueness of far field expansions is important for two reasons. First, redundant coefficients could mean that a small value is represented as the difference of two large values leading to numerical instability. Second, if the far field expansion of a fixed function, $s(x) = \sum_{i=1}^{N} \Phi(x - t_i)$, is unique then it is often possible to shift the centre of a truncated expansion indirectly without using any knowledge of the underlying centres and weights. The advantage of such indirect shifting over direct series formation is a flop count which depends only on the number of terms in the expansion, and not on the number of centres in the cluster. This can result in significantly faster code. Furthermore, since the uniqueness implies the indirectly obtained series is identical with that which would have been obtained directly, the indirectly obtained series enjoys the same error bound as the directly obtained one.

We will now prove a general uniqueness lemma from which uniqueness of series expansions of the form (3.17) follows as a special case. Recall that a function g defined for all x in some subset $D \subset \mathbb{R}^n$ is said to be homogeneous of degree γ on D if

$$g(\lambda x) = \lambda^\gamma g(x)$$

for all $\lambda > 0$ and $x \in \mathbb{R}^n$ such that both x and $\lambda x \in D$. (Some authors use the term positively homogeneous of degree γ for this property).

Lemma 3.10. *Suppose $\gamma, R \in \mathbb{R}$ and that a function $f : D \subset \mathbb{R}^n \to \mathbb{R}$ can be expanded in two ways*

$$\sum_{\ell=0}^{\infty} U_\ell(x) = f(x) = \sum_{\ell=0}^{\infty} V_\ell(x),$$

both series converging absolutely and uniformly to $f(x)$ for all $|x| \geq R$, where for each ℓ , U_ℓ and V_ℓ are continuous homogeneous functions of degree $\gamma - \ell$. Then for each ℓ, $U_\ell(x) = V_\ell(x)$ for all $|x| \geq R$.

Proof. Since the absolute series converge uniformly on $|x| = R$ there exists an $M < \infty$ such that

$$\max_{|x|=R} \{\max\{|U_\ell(x)|, \ |V_\ell(x)|\}\} \leq M,$$

for all $\ell \in \mathbb{N}_0$. Hence, using the homogeneity,

$$\max\{|U_\ell(x)|, \ |V_\ell(x)|\} \leq M|x|^{\gamma-\ell}/R^{\gamma-\ell}, \tag{3.18}$$

for all x such that $|x| \geq R$, and all $\ell \in \mathbb{N}_0$.

Now suppose U_ℓ and V_ℓ differ for some ℓ's. Let j be the first index for which they differ. Then for all $|x| \geq R$

$$\begin{aligned} 0 &= \left(\frac{|x|}{R}\right)^{j-\gamma} \{f(x) - f(x)\} \\ &= \left(\frac{|x|}{R}\right)^{j-\gamma} \{U_j(x) - V_j(x)\} + \sum_{\ell>j} \left(\frac{|x|}{R}\right)^{j-\gamma} \{U_\ell(x) - V_\ell(x)\}. \end{aligned} \tag{3.19}$$

But from (3.18)

$$\begin{aligned} \left|\sum_{\ell>j} \left(\frac{|x|}{R}\right)^{j-\gamma} \{U_\ell(x) - V_\ell(x)\}\right| &\leq 2M \sum_{\ell>j} \left(\frac{|x|}{R}\right)^{j-\ell} \\ &= o(1) \qquad \text{as } |x| \to \infty. \end{aligned}$$

Hence from (3.19)

$$|U_j(x) - V_j(x)| = o(|x|^{\gamma-j}) \quad \text{as } |x| \to \infty.$$

Since $U_j - V_j$ is homogeneous of degree $\gamma - j$ on D this implies that it is identically zero on D. ∎

3.4 Efficient formation of the far field series

In the previous sections we have developed far field expansions with the intention of using them for fast evaluation of generalised multiquadric RBF's. In order that these expansions be suitable for this task they must be inexpensive both to form and to evaluate. The purpose of this section is to show that the expansions can be formed in an efficient recursive manner.

Given a single centre $t \in \mathbb{R}^n$, with unit weight, the corresponding truncated expansion of Section 3.2 is

$$\Phi(x-t) = \left((x-t)^2 + \tau^2\right)^{k/2} = \sum_{\ell=0}^{\infty} P_\ell^{(k)}(t^2+\tau^2, -2\langle t, x\rangle, x^2)/|x|^{2\ell-k} \qquad (3.20)$$

Writing $G_\ell(x) = P_\ell^{(k)}(t^2+\tau^2, -2\langle t, x\rangle, x^2)$, G_ℓ is a homogeneous polynomial of degree ℓ in x, with coefficients depending on k, τ and t. The expansion for a single centre, with corresponding weight d, then becomes

$$\sum_{\ell=0}^{p+k} dG_\ell(x)/|x|^{2\ell-k}. \qquad (3.21)$$

The expansion of a cluster is formed by summing the expansions (3.21) corresponding to each centre, and has the form

$$\sum_{\ell=0}^{p+k} Q_\ell(x)/|x|^{2\ell-k}, \qquad (3.22)$$

where each Q_ℓ is a homogeneous polynomial of degree ℓ. Lemma 3.5 implies that the polynomials G_ℓ satisfy the three term recurrence

$$G_\ell(x) = \begin{cases} 1, & \ell = 0, \\ -k\langle x, t\rangle, & \ell = 1, \\ A_\ell\langle x, t\rangle G_{\ell-1}(x) + B_\ell\, x^2(t^2+\tau^2)\, G_{\ell-2}(x), & \ell \geq 2, \end{cases} \qquad (3.23)$$

where

$$A_\ell = -2\frac{k/2-\ell+1}{\ell}, \qquad B_\ell = -\frac{\ell-k-2}{\ell}.$$

The recurrence is very simple to implement as is demonstrated by the code fragment in Algorithm 3.1 on page 107 for the special case of 2-dimensions. The code fragment employs the notation of Example 3.8.

Recall that the $\binom{\ell+n-1}{\ell}$ monomials of exact degree ℓ, $\{x^\alpha : |\alpha| = \ell\}$, form a basis for the homogeneous polynomials of degree ℓ on $\mathbb{R}^n$. Represent the polynomials G_ℓ in terms of these monomials, *i.e.*, let

$$G_\ell(x) = \sum_{|\alpha|=\ell} a_\alpha^\ell x^\alpha,$$

for some coefficients a_α^ℓ. Then, from the recurrence (3.23), each coefficient of G_ℓ can be calculated using at most n coefficients of $G_{\ell-1}$ and at most n coefficients of $G_{\ell-2}$. Specifically, if e_i is the multiindex with 1 in the ith position and 0 elsewhere then the recurrence (3.23) implies,

$$a_\alpha^\ell = A_\ell \sum_{i=1}^{n} t_i a_{\alpha-e_i}^{\ell-1} + B_\ell(t^2+\tau^2)\sum_{i=1}^{n}(a_{\alpha-2e_i}^{\ell-2})^2,$$

where a_α^ℓ is taken to be zero if any component of α is negative. It follows that all the numerator polynomials $\{Q_\ell\}_{\ell=0}^{p+k}$ in the truncated expansion (3.22) of an m centre cluster can be formed (that is their $\binom{n+p+k}{p+k}$ coefficients calculated) in $\mathcal{O}\big(mn\binom{n+p+k}{p+k}\big)$ floating point operations. This quantity is $\mathcal{O}\big(mn(p+k)^n\big)$ when the dimension n is less than the degree $p+k$.

3.5 A subspace of polynomials

In this section we will investigate a subspace of polynomials in n variables. This space will arise in Section 3.6 and the aim of that section will be to translate a member of this subspace. It will shown that, modulo a low degree polynomial, this subspace is closed under translation of the underlying Cartesian coordinate system.

Throughout this section and the next n will be fixed and any complexity estimates will be expressed as a function of polynomial degree only. Thus a typical estimate might take the form $\mathcal{O}((p+k)^n)$. In such expressions multiplicative order constants depending on n have been suppressed, and we will be interested in the estimate only when the argument $p+k$ is bigger than n.

The following standard spaces will be used.

- π_j^n Polynomials of total degree not exceeding j in n variables.
- $\mathcal{H}_j^n$ homogeneous polynomials of degree j in n variables.

Also, for given function spaces S and T, define new spaces as follows.

$$\begin{aligned} ST &= \{s(\cdot)\,t(\cdot) \ : \ s\in S,\ t\in T\}, & & \\ S\oplus T &= \{s(\cdot)+t(\cdot) \ : \ s\in S,\ t\in T\}, & & S\cap T=\{0\}, \\ sT &= \{s(\cdot)\,t(\cdot) \ : \ t\in T\}, & & s\in S. \end{aligned}$$

The subspaces of polynomials that are the subject of this section are defined by

$$S_j^n = \left\{ q\in\pi_{2j}^n : q(\cdot)=\sum_{\ell=0}^{j} q_\ell(\cdot)|\cdot|^{2(j-\ell)}, q_\ell\in\mathcal{H}_\ell^n \right\}. \tag{3.24}$$

Apart from 0, the polynomials of S_j^n have total degree no greater than $2j$ and no less than j. It follows from Lemma 3.10 that $q \in S_n^j$ is uniquely determined by the homogeneous polynomials $\{q_\ell\}_{\ell=0}^j$ and thus by the coefficients of those polynomials with respect to some appropriate basis. Hence

$$\dim S_j^n = \sum_{\ell=0}^{j} \dim H_\ell^n. \tag{3.25}$$

Theorem 3.11. *S_j^n in invariant under orthogonal transformation of the underlying coordinate system,* i.e., *if $q \in S_j^n$ then $q(Q\,\cdot) \in S_j^n$ for orthogonal Q.*

Proof. For each i, the component function $f_i(x) = (Qx)_i$ is homogeneous in x of exact degree 1 since $x \mapsto Qx$ is a linear operation. Thus

$$(Qx)^\alpha = (Qx)_1^{\alpha_1}(Qx)_2^{\alpha_s}\ldots(Qx)_n^{\alpha_n}$$

is homogeneous of exact degree $|\alpha|$. It follows that $q_\ell(Q\,\cdot)$ is homogeneous of degree ℓ if q_ℓ is. Finally, since Q is orthogonal,

$$|Q\cdot| = |\cdot|,$$

and the result follows. ∎

Before we prove translation invariance of S_j^n we will make a few simple observations regarding these spaces.

Lemma 3.12. *The spaces S_j^n satisfy the following relations.*

(i). $S_{j+1}^n = \left(|\cdot|^2\, S_j^n\right) \oplus \mathcal{H}_{j+1}^n,$

(ii). $\mathcal{H}_1^n\, S_j^n \subset S_{j+1}^n,$

(iii). $S_j^n \subset S_{j+1}^n \oplus \mathcal{H}_j^n.$

Proof. Let $q \in S_{j+1}^n$ and let $\{q_\ell\}_{\ell=0}^{j+1}$ be the polynomials such that

$$q = \sum_{\ell=0}^{j+1} |\cdot|^{2(j+1-\ell)} q_\ell, \qquad q_\ell \in \mathcal{H}_\ell^n.$$

The observation that

$$q = |\cdot|^2 h + q_{j+1}$$

where

$$h = \left(\sum_{\ell=0}^{j} |\cdot|^{2(j-\ell)} q_\ell\right) \in S_j^n$$

proves part (i).

Now let $p \in \mathcal{H}_1^n$. Then for each ℓ, $0 \le \ell \le j$, the product $\tilde{q}_{\ell+1} = pq_\ell \in \mathcal{H}_{\ell+1}^n$. Thus

$$pq = \sum_{\ell=0}^{j} |\cdot|^{2(j-\ell)} \tilde{q}_{\ell+1} = \sum_{k=1}^{j+1} |\cdot|^{2(j+1-k)} \tilde{q}_k \in S_{j+1}^n,$$

which shows part(ii).

Part (iii) follows from part (i) with j replace by $j-1$. ∎

Theorem 3.13. *S_j^n is translation invariant modulo polynomials of degree $j-1$,* i.e., *for any $q \in S_j^n$ and $u \in \mathbb{R}^n$, $q(\cdot - u) \in S_j^n \oplus \pi_{j-1}^n$.*

Proof. The proof is by induction on j. The result is trivially true in the case $j = 0$ since S_0^n is the space of constants and π_{-1}^n is the singleton $\{0\}$.

Now assume the result for $k = 0, 1, 2, \ldots, j$, let $q \in S_{j+1}^n$ and let $u \in \mathbb{R}^n$. Then by Lemma 3.12, part (i),

$$q(x-u) = |x-u|^2\, h(x-u) + q_{j+1}(x-u) \tag{3.26}$$

where $h \in S_j^n$ and $q_{j+1} \in \mathcal{H}_{j+1}^n$. By the induction hypothesis, $h(\cdot - u) \in S_j^n \oplus \pi_{j-1}^n$. Thus

$$h(x-u) = \tilde{h}_j(x) + \tilde{h}_{j-1}(x) + \tilde{h}_<(x) \tag{3.27}$$

where $\tilde{h}_j \in S_j^n$, $\tilde{h}_{j-1} \in \mathcal{H}_{j-1}^n$ and $\tilde{h}_< \in \pi_{j-2}^n$. Since $q_{j+1} \in \mathcal{H}_{j+1}^n$,

$$q_{j+1}(x-u) = q_{j+1}(x) - \tilde{q}_<(x) \tag{3.28}$$

where $\tilde{q}_<(x) \in \pi_j^n$. Expand (3.26) to get

$$q(x-u) = \left(|x|^2 - 2\langle x, u\rangle + |u|^2\right)\left(\tilde{h}_j(x) + \tilde{h}_{j-1} + \tilde{h}_<\right) + q_{j+1}(x) + \tilde{q}_<(x). \tag{3.29}$$

Consider each term of in the expansion of this product:

$$\begin{aligned}
|\cdot|^2 \tilde{h}_j &\in S_{j+1}^n && \text{by Lemma 3.12, part (i)}\\
-2\langle \cdot, u\rangle \tilde{h}_j &\in S_{j+1}^n && \text{by Lemma 3.12, part (ii)}\\
|u|^2 \tilde{h}_j &\in S_j^n \subset S_{j+1}^n \oplus \mathcal{H}_j^n && \text{by Lemma 3.12, part (iii)}\\
|\cdot|^2 \tilde{h}_{j-1} &\in S_{j+1}^n && \text{by definition of } S_{j+1}^n\\
-2\langle \cdot, u\rangle \tilde{h}_{j-1} &\in \mathcal{H}_j^n &&\\
|u|^2 \tilde{h}_{j-1} + |\cdot - u|^2 \tilde{h}_< &\in \pi_j^n. &&
\end{aligned}$$

Thus it follows that $q(\cdot - u) \in S_{j+1}^n \oplus \pi_j^n$. The result follows by induction. ∎

In computations, a polynomial $p \in S_j^n$ may be known in terms of the monomial basis, but what is actually required are the polynomials $\{q_\ell\}_{\ell=0}^{j}$ such that

$$p(x) = \sum_{\ell=0}^{j} q_\ell(x)|x|^{2(j-\ell)}. \tag{3.30}$$

Since the polynomials $\{q_\ell\}$ are homogeneous, for a given ℓ, q_ℓ must be determined entirely by those terms of p that are homogeneous of degree $2j-\ell$. Thus the problem of determining $\{q_\ell\}$ may be broken down into homogeneous parts. Hence, without loss of generality, assume that p is a given homogeneous polynomial of degree $\ell + 2k$ such that

$$p(x) = |x|^{2k} q(x) \tag{3.31}$$

with q unknown and to be determined from p. Since

$$p(x) = |x|^{2k} q(x) = |x|^2 \left(|x|^{2(k-1)} q(x)\right),$$

if q can be determined in the case where $k = 1$, the more general problem may be solved in an inductive manner.

Let $\{p_j\}_{j=0}^{\ell+2}$ and $\{q_i\}_{i=0}^{\ell}$ be homogeneous polynomials in $x_2, \ldots, x_n$ such that

$$p(x) = \sum_{j=0}^{\ell+2} x_1^{\ell+2-j} p_j(\bar{x}), \quad q(x) = \sum_{i=0}^{\ell} x_1^{\ell-i} q_i(\bar{x}), \quad \text{and} \quad p(x) = |x|^2 q(x),$$

where, if $x = (x_1, \ldots, x_n)$ then $\bar{x} = (x_2, \ldots, x_n)$. Using this same notation,

$$|x|^2 = x_1^2 + |\bar{x}|^2$$

and hence

$$\begin{aligned}
\sum_{j=0}^{\ell+2} x_1^{\ell+2-j} p_j(\bar{x}) &= \left(x_1^2 + |\bar{x}|^2\right) \sum_{i=0}^{\ell} x_1^{\ell-i} q_i(\bar{x}) \\
&= x_1^{\ell+2} q_0(\bar{x}) + x_1^{\ell+1} q_1(\bar{x}) + \left\{ \sum_{i=2}^{\ell} x_1^{\ell+2-i} \left(q_i(\bar{x}) + |\bar{x}|^2 q_{i-2}(\bar{x})\right) \right\} \\
&\qquad + x_1 |\bar{x}|^2 q_{\ell-1}(\bar{x}) + |\bar{x}|^2 q_\ell(\bar{x}).
\end{aligned}$$

Equating coefficients the polynomials q_i may now be written in terms of the polynomials p_j.

$$q_0(\bar{x}) = p_0(\bar{x}),$$
$$q_1(\bar{x}) = p_1(\bar{x}),$$

$$\begin{aligned}
q_2(\bar{x}) &= p_2(\bar{x}) - |\bar{x}|^2 q_0(\bar{x}),\\
q_3(\bar{x}) &= p_3(\bar{x}) - |\bar{x}|^2 q_1(\bar{x}),\\
&\vdots\\
q_{\ell-1}(\bar{x}) &= p_{\ell-1}(\bar{x}) - |\bar{x}|^2 q_{\ell-3}(\bar{x}),\\
q_\ell(\bar{x}) &= p_\ell(\bar{x}) - |\bar{x}|^2 q_{\ell-2}(\bar{x}).
\end{aligned}$$

Multiplication of a polynomial by a monomial corresponds to a relabelling of coefficients and computationally corresponds to assignment or addition. Since,

$$|\bar{x}|^2 = x_2^2 + \cdots + x_n^2$$

is just the sum of $n-1$ monomials, for fixed i the product $|\cdot|^2 q_i(\cdot)$ may calculated with $\mathcal{O}(nC_i)$ additions, where $C_i = \dim \mathcal{H}_i^{n-1}$. It is well known that

$$\dim \mathcal{H}_i^{n-1} = \binom{i+n-2}{n-2} = \frac{(i+n-2)!}{i!(n-2)!} = \frac{1}{(n-2)!}\big((i+n-2)\cdots(i+1)\big) = \mathcal{O}(i^{n-2}),$$

and hence $|\cdot|^2 q_i(\cdot)$ may be calculated in $\mathcal{O}(i^{n-2})$ operations. It now follows that all of the polynomials $\{q_i\}_{i=0}^{\ell}$ may be calculated in $\mathcal{O}(\ell^{n-1})$ operations.

Since the more general problem of (3.31) may be solved by k applications of this simpler case, $q(x) = p(x)/|x|^{2k}$ may be calculated in

$$\sum_{i=0}^{k-1} \mathcal{O}\big((\ell+2i)^{n-1}\big) = \mathcal{O}\big((\ell+2k)^n\big)$$

operations. Applying this to each homogeneous part of (3.30) gives the following lemma.

Lemma 3.14. *Let $n \in \mathbb{N}$. There exists a constant C depending only on n with the following property. Given any polynomial $p \in S_j^n$ the polynomials $\{q_\ell\}_{\ell=0}^{j}$ such that $q_\ell \in \mathcal{H}_\ell^n$ and*

$$p = \sum_{\ell=0}^{j} |\cdot|^{2(j-\ell)} q_\ell\ ,$$

may be determined in no more than Cj^{n+1} operations.

3.6 Translation of a far field expansion

The uniqueness of the far field expansions makes it is possible to shift the centre of a truncated expansion knowing only its coefficients, and without any direct knowledge of the underlying centres and weights. As the operation count for indirect translation depends on the length of the series, not the number of centres, indirect translation can be significantly faster than direct formation of series for clusters with many centres.

The precise problem we address is the following. Let

$$s_p(x) = \sum_{\ell=0}^{p+k} Q_\ell(y)/|y|^{2\ell-k}, \qquad y = x - u \neq 0, \tag{3.32}$$

where Q_ℓ are homogeneous polynomials of degree ℓ, be an expansion similar to (3.17) or (3.22), but centred at $u \neq 0$ rather than 0. We wish to shift the centre of expansion to the origin. That is we seek homogenous polynomials $\{\widehat{Q}_\ell\}$, $\widehat{Q}_\ell$ being of degree ℓ, so that

$$s_p(x) = \sum_{\ell=0}^{p+k} \widehat{Q}_\ell(x)/|x|^{2\ell-k} + \mathcal{O}(1/|x|^{p+1}), \tag{3.33}$$

as $|x| \to \infty$. We will show that translations of truncated expansions of the form (3.32) into expansions of the form (3.33) may be performed in $\mathcal{O}\left((p+k)^{n+1}\right)$ operations using simple polynomial manipulations.

3.6.1 The cost of multiplication

In this subsection it will be shown that the product of two homogeneous polynomials of degree ℓ in n variables may be computed in $\mathcal{O}(\ell^{n-1}\log\ell)$ operations.

Let p be a homogeneous polynomial of degree ℓ. Since p is homogeneous,

$$p(x) = p(x_1, x_2, \ldots, x_n) = x_n^\ell p\left(\frac{x_1}{x_n}, \frac{x_2}{x_n}, \ldots, \frac{x_{n-1}}{x_n}, 1\right), \qquad x_n \neq 0.$$

Furthermore given $x_n^\ell p(\ldots)$ for all x with $x_n \neq 0$, $p(x)$ can be recovered on the hyperplane $x_n = 0$ by continuity. Thus for the purposes of the multiplication and division that are the subject of this section, we may consider multiplication and division of general, that is probably inhomogeneous, polynomials of degree ℓ in $n-1$ variables rather than of homogeneous polynomials of degree ℓ in n variables.

Let p and q be two polynomials of degree ℓ in $n-1$ variables. Then their product is

$$p(x)q(x) = \left(\sum_{|\alpha|\le\ell} a_\alpha x^\alpha\right)\left(\sum_{|\beta|\le\ell} b_\beta x^\beta\right) = \sum_{|\alpha|\le 2\ell}\left(\sum_{0\le\beta\le\alpha} a_\beta b_{\alpha-\beta}\right) x^\alpha,$$

the Cauchy product. The convolution producing the coefficients of the product can be computed in $\mathcal{O}(\ell^{n-1}\log\ell)$ operations by FFTs. It now follows that the homogeneous polynomial multiplication above can also be carried out in $\mathcal{O}(\ell^{n-1}\log\ell)$ operations.

3.6.2 Translation by convolution

In this subsection it will be shown that translation of the far field series may be performed by convolution.

Throughout this subsection when we speak of forming a polynomial we mean finding its coefficients with respect to a basis, usually the monomial basis. When we speak of forming a truncated expansion of the type (3.22), we mean finding the coefficients of all the relevant numerator polynomials.

First we set

$$Q(y) = \sum_{\ell=0}^{p+k} Q_\ell(y)|y|^{2(p+k-\ell)}. \tag{3.34}$$

Then

$$s_p(x) = Q(y)/|y|^{2p+k}, \qquad y = x - u \neq 0. \tag{3.35}$$

Since we already have all of the Q_ℓ, all we need to do to form Q is form the polynomials $|\cdot|^{2(p+k-\ell)}$ and then form the products $Q_\ell(\cdot)|\cdot|^{2(p+k-\ell)}$. Form $|\cdot|^{2j}$, $j = 0, \dots, p+k$ once and store. Each $|\cdot|^{2j-2}$ is homogeneous of degree $2j-2$ and therefore involves $\mathcal{O}(j^{n-1})$ coefficients. The polynomial $|\cdot|^{2j}$ may be obtained from $|\cdot|^{2j-2}$ with n additions for each coefficient in $|\cdot|^{2j-2}$. Hence the cost of forming the $|\cdot|^{2j}$'s is $\mathcal{O}((p+k)^n)$ operations. Each of the products $Q_\ell(\cdot)|\cdot|^{2(p+k-\ell)}$ is the product of two homogeneous polynomials and is of degree no greater than $2(p+k)$. Hence we can calculate each product in $\mathcal{O}\big((p+k)^{n-1}\log(p+k)\big)$ operations. As there are $p+k+1$ of these products in Q, forming Q takes $\mathcal{O}\big((p+k)^n \log(p+k)\big)$ operations.

We proceed to shift the centre of expansion of Q by setting

$$\widetilde{Q}(x) = Q(x-u), \qquad x \in \mathcal{R}^n. \tag{3.36}$$

A translation of this sort can be done simply and quickly by convolution. For example, using the scaled monomial basis $V_\alpha(x) = x^\alpha/\alpha!$ (α a multi-index), we have

$$\begin{aligned}
p(x-u) &= \sum_{|\alpha|<k} a_\alpha \, V_\alpha(x-u) \\
&= \sum_{|\alpha|<k} a_\alpha \frac{(x-u)^\alpha}{\alpha!} \\
&= \sum_{|\alpha|<k} \frac{a_\alpha}{\alpha!} \sum_{\beta<\alpha} \binom{\alpha}{\beta} x^\beta (-u)^{(\alpha-\beta)} \\
&= \sum_{|\alpha|<k} a_\alpha \sum_{\beta<\alpha} \frac{x^\beta}{\beta!} \frac{(-u)^{(\alpha-\beta)}}{(\alpha-\beta)!}
\end{aligned}$$

$$= \sum_{|\beta|<k} \frac{x^\beta}{\beta!} \sum_{\alpha<\beta} a_\alpha \frac{(-u)^{(\alpha-\beta)}}{(\alpha-\beta)!}.$$

Thus an n-dimensional convolution of $\{a_\alpha\}$ and $\{(-u)^\alpha/\alpha!\}$ gives the coefficients of the translated polynomial. Again this can be computed in $\mathcal{O}\big((p+k)^n \log(p+k)\big)$ operations by an FFT method. This gives us $\widetilde{Q}$ in terms of the monomial or scaled monomial basis.

The next task is to recast $\widetilde{Q}$ into a sum of products of powers of $|x|$ and homogeneous polynomials. By Theorem 3.13 we know that

$$\widetilde{Q}(x) = \sum_{\ell=0}^{p+k} q_\ell(x)|x|^{2(p+k-\ell)} + q_{\text{low}}(x) \tag{3.37}$$

where the q_ℓ are homogeneous of degree ℓ and q_{low} is some polynomial of degree $p+k-1$ or less. By Lemma 3.14, these homogeneous polynomials q_ℓ can be calculated from $\widetilde{Q}$ in $\mathcal{O}\big((p+k)^{n+1}\big)$ operations.

Combining equations (3.35) and (3.36) and appealing to Lemma 3.6 gives

$$\begin{aligned}
s_p(x) &= Q(x-u)/|x-u|^{2p+k} \\
&= \widetilde{Q}(x)/|x-u|^{2p+k} \qquad\qquad (3.38)\\
&= \widetilde{Q}(x) \sum_{m=0}^{\infty} P_m^{(-2p-k)}(u^2, -2\langle x,u\rangle, x^2)/|x|^{2p+k+2m} \\
&= \left(\sum_{\ell=0}^{p+k} q_\ell(x)|x|^{2(p+k-\ell)} + q_{\text{low}}(x) \right) \\
&\quad \times \left(\sum_{m=0}^{\infty} P_m^{(-2p-k)}(u^2, -2\langle x,u\rangle, x^2)/|x|^{2p+k+2m} \right) \\
&= \sum_{\ell=0}^{p+k} \sum_{m=0}^{\infty} q_\ell(x) P_m^{(-2p-k)}(u^2, -2\langle x,u\rangle, x^2)/|x|^{2(m+\ell)-k} + \mathcal{O}(1/|x|^{p+1}) \\
&= \sum_{\ell=0}^{p+k} \left(\sum_{j=0}^{\ell} q_j(x) P_{\ell-j}^{(-2p-k)}(u^2, -2\langle x,u\rangle, x^2) \right) /|x|^{2\ell-k} + \mathcal{O}(1/|x|^{p+1}) \\
&= \sum_{\ell=0}^{p+k} \widehat{Q}_\ell(x)/|x|^{2\ell-k} + \mathcal{O}(1/|x|^{p+1}).
\end{aligned}$$

The sums of products

$$\widehat{Q}_\ell(x) = \sum_{j=0}^{\ell} q_j(x) P_{\ell-j}^{(-2p-k)}(u^2, -2\langle x,u\rangle, x^2), \quad 0 \le \ell \le p+k, \tag{3.39}$$

can be computed simultaneously as homogeneous parts of the product

$$\left[\sum_{j=0}^{p+k} q_j(\cdot)\right]\left[\sum_{m=0}^{p+k} P_m^{-2p+k}\left(u^2, -2\langle\cdot, u\rangle, (\cdot)^2\right)\right].$$

Hence they can be computed by a single FFT convolution in $\mathcal{O}\left((p+k)^n \log(p+k)\right)$ operations.

3.7 Conversion to a near field series

The final step in the process of forming expansions for the FMM is to convert the far field series into a near field, or Taylor, series. At the implementation level, this step is almost identical to the first part of the translation of the far field series.

Define two non-intersecting discs:

$$D_{\text{eval}} = \{x : |x| \le r\},$$
$$D_{\text{src}} = \{x : |x-u| \le \sqrt{(\theta r)^2 - \tau^2}\}, \qquad \theta > 0.$$

Let

$$s_p(x) = \sum_{\ell=0}^{p+k} Q_\ell(y)/|y|^{2\ell-k}, \qquad y = x-u \ne 0,$$

be a far field series, such as (3.17) or (3.32), of $s(x) = \sum_{i=1}^N d_i \Phi(x - t_i)$ due to a cluster of centres $\{t_i\}$ located inside D_{src}. Then by Theorem 3.9, s_p approximates s well on D_{eval}. We wish find to a near field series that approximates s_p, and thus s, on D_{eval}.

Proceeding in an identical fashion to Section 3.6.2, we see that we may calculate the polynomial $\widetilde{Q}$ such that

$$s_p(x) = \widetilde{Q}(x)/|x-u|^{2p+k}$$

in $\mathcal{O}\left((p+k)^n \log(p+k)\right)$ operations. When translating the far field expansion to another far field expansion, we essentially convolved $\widetilde{Q}$ with the far field series for $|\cdot - u|^{-(2p+k)}$. To get the near field, all we need do is convolve $\widetilde{Q}$ with the near field series for $|\cdot - u|^{-(2p+k)}$.

The next result gives an explicit expression for the Maclaurin series of $\Phi(\cdot - u) = \left((\cdot - u)^2 + \tau^2\right)^{k/2}$ together with an estimate of the error in approximation by truncating this series. Specialising to the case $\tau = 0$ in this lemma gives the Maclaurin series for $|\cdot - u|^k$.

Lemma 3.15. *Let $k \in \mathbb{Z}$ be odd, and $u \in \mathbb{R}^n \setminus \{0\}$ and $\tau \ge 0$. For all $x \in \mathbb{R}^n$ with $|x| < \sqrt{u^2 + \tau^2}$,*

$$\Phi(x-u) = \left((x-u)^2 + \tau^2\right)^{k/2} = \sum_{\ell=0}^{\infty} P_\ell^{(k)}(x^2, -2\langle u, x\rangle, u^2 + \tau^2)/\left(\sqrt{u^2+\tau^2}\right)^{2\ell-k} \qquad (3.40)$$

where the polynomials $P_\ell^{(k)}$ *are defined in Equation* (3.4). *Moreover,*

$$T_q\big(\Phi(\cdot - u)\big)(x) := \sum_{\ell=0}^{q} P_\ell^{(k)}(x^2, -2\langle u, x\rangle, u^2+\tau^2)/\big(\sqrt{u^2+\tau^2}\big)^{2\ell-k}, \tag{3.41}$$

is the Maclaurin polynomial of degree q *of* $\Phi(\cdot - u)$. *When* $|x| < \sqrt{u^2+\tau^2}$ *and* $q \in \mathbb{N}$,

$$\left|\Phi(x-u) - \sum_{\ell=0}^{q} P_\ell^{(k)}(x^2, -2\langle u, x\rangle, u^2+\tau^2)/\big(\sqrt{u^2+\tau^2}\big)^{2\ell-k}\right|$$
$$\leq \begin{cases} \big(\sqrt{u^2+\tau^2}\big)^k \left(\dfrac{|x|}{\sqrt{u^2+\tau^2}}\right)^{q+1} \dfrac{\sqrt{u^2+\tau^2}}{\sqrt{u^2+\tau^2}-|x|}, & \textit{if } k > 0, \\ \dbinom{q-k}{q+1} \big(\sqrt{u^2+\tau^2}\big)^k \left(\dfrac{|x|}{\sqrt{u^2+\tau^2}}\right)^{q+1} \left(\dfrac{\sqrt{u^2+\tau^2}}{\sqrt{u^2+\tau^2}-|x|}\right)^{-k}, & \textit{if } k < 0. \end{cases} \tag{3.42}$$

Proof. Assume firstly that $x \neq 0$. Let $a = x^2$, $b = -2\langle u, x\rangle$ and $c = u^2 + \tau^2$. Then

$$\Phi(x-u) = \big(x^2 - 2\langle u, x\rangle + u^2 + \tau^2\big)^{k/2} = f_k(1),$$

where f_k is the function that is defined in (3.5). Since $a, c > 0$, $b^2 \leq 4ac$, and $1 = |z| < \sqrt{c/a} = \sqrt{u^2+\tau^2}/|x|$, Lemma 3.4 may be applied with $\nu = q$ to yield Equations (3.40) and (3.42) when $x \neq 0$. The results for $x = 0$ follow by continuity.

It remains to show that $T_q\big(\Phi(\cdot - u)\big)$ is the Maclaurin polynomial of $\Phi(\cdot - u)$. Observe from (3.4) that

$$P_\ell^{(k)}(a, b, c) = P_\ell^{(k)}(x^2, -2\langle u, x\rangle, u^2+\tau^2)$$

is either a homogeneous polynomial of exact degree ℓ in x, or is trivial. Hence, by Equation (3.42), $T_q\big(\Phi(\cdot - u)\big)$ is a polynomial of total degree q in x such that

$$\Big|\Phi(x-u) - T_q\big(\Phi(\cdot - u)\big)(x)\Big| = \mathcal{O}|x|^{q+1} \text{ as } |x| \to 0.$$

The result follows since the only such polynomial is the Maclaurin polynomial. ∎

3.8 Numerical results

In this section we present numerical results generated by an initial, non-optimised, implementation of a hierarchical evaluator for generalised multiquadrics. Note that local expansions were not used in this code.

The current implementation is based on a hierarchical subdivision of an initial box containing all the centres using a binary tree of panels. Associated with a panel are the centres

N	Direct time	Algorithm time	Ratio
1,000	3.20 (-1)	1.30 (-1)	2.46
2,000	1.312 (0)	3.30 (-1)	3.98
4,000	5.358 (0)	7.91 (-1)	6.77
8,000	2.745 (1)	1.762 (0)	15.58
16,000	1.098 (2)	3.665 (0)	29.96
32,000	4.394 (2)	8.382 (0)	52.42

Table 3.1: Results of numerical experiments with a generalised multiquadric fast evaluator.

lying within it, a far field expansion, and a distance from the panel's midpoint at which the far field expansion approximates the influence of the panel to sufficient accuracy. Panels are divided generating children if they contain more than a critical number of centres.

Pseudo code for recursive and non-recursive evaluators appropriate for use with such a binary tree evaluation structure is sketched in [9, pp. 8–11]. Nominally the discussion there is limited to an $\mathbb{R}^1$, rather than $\mathbb{R}^n$, setting but the generalisation is immediate.

Table 3.1 above gives times in seconds on an Intel Celeron based machine for various evaluation tasks in $\mathbb{R}^2$. An entry of the form $d_0.d_1d_2d_3(e)$ in the table with d_0, d_1, d_2, d_3 decimal digits represents the number $d_0.d_1d_2d_3 \times 10^e$. In the numerical experiments the centres are uniformly distributed on $[0,1]^2$, the multiquadric parameter τ is taken as $1/\sqrt{N}$, where N is the number of centres, and ϕ is the ordinary multiquadric $\phi(r) = \sqrt{r^2+\tau^2}$. All the coefficients d_i were taken as 1 and the task was to evaluate the spline at the centres to with an infinity norm relative accuracy of 10^{-6}. The code used was structured as a general evaluator and the symmetry inherent in this matrix-vector product test problem was not exploited.

It can be seen from the table that even this initial, non-optimised, implementation is substantially faster than direct evaluation. Thus the methods described here will allow use of multiquadric RBF's in much bigger problems than previously possible. We would expect even better performance as the code is developed to incorporate such features as conversion of far field to local expansions.

Input: A centre $t \in \mathbb{R}^2$, the corresponding weight d, the generalised multiquadric parameters k and τ, and the desired order of expansion p.
Output: The coefficients $G(\ell, j)$ of the homogeneous numerator polynomials in the expansion of this single centre. On output $G(\ell, j)$ is the coefficient of $x_1^{\ell-j} x_2^j$ in the homogeneous polynomial dG_ℓ of Equation (3.21).
PolynomialGenerator(t, d, k, τ, p)

$G(0,0) = d$, $G(1,0) = -d * k * t_1$, $G(1,1) = -d * k * t_2$
for $\ell = 2$ **to** $p + k$
    $a = A_\ell$, $b = B_\ell * (|t|^2 + \tau^2)$
    $\text{tmp} = a * G(\ell - 1, 0)$
    $G(\ell, 0) = \text{tmp} * t_1$
    $G(\ell, 1) = \text{tmp} * t_2$
    for $j = 0$ **to** $\ell - 2$
        $\text{tmp} = b * G(\ell - 2, j)$
        $G(\ell, j) = G(\ell, j) + \text{tmp}$
        $G(\ell, j+2) = \text{tmp}$
        $\text{tmp} = a * G(\ell - 1, j + 1)$
        $G(\ell, j+1) = G(\ell, j+1) + \text{tmp} * t_1$
        $G(\ell, j+2) = G(\ell, j+2) + \text{tmp} * t_2$
    end
end

Algorithm 3.1: Code fragment to generate the numerator polynomial coefficients in the expansion of a generalised multiquadric in 2-dimensions.

Chapter 4

Fitting a surface to a cloud of points

In this chapter we outline an implicit surface approach to the problem of fitting a surface to a cloud of points. Our approach is based on interpolation with a polyharmonic radial basis function, followed by isosurface extraction. We demonstrate that the method is successful on objects of varying geometry, and that it can handle large clouds with in excess of 70,000 points. The function used to implicitly define the surface is a Radial Basis Function (RBF) with at least one centre for each point in the input cloud. As was demonstrated in Chapter 1, techniques such as the Fast Multipole Method (FMM) are absolutely essential if problems of this size are to be tackled with RBFs. Some of the material in this chapter is also discussed in [6]. That paper contains a more comprehensive discussion of some of the computational issues.

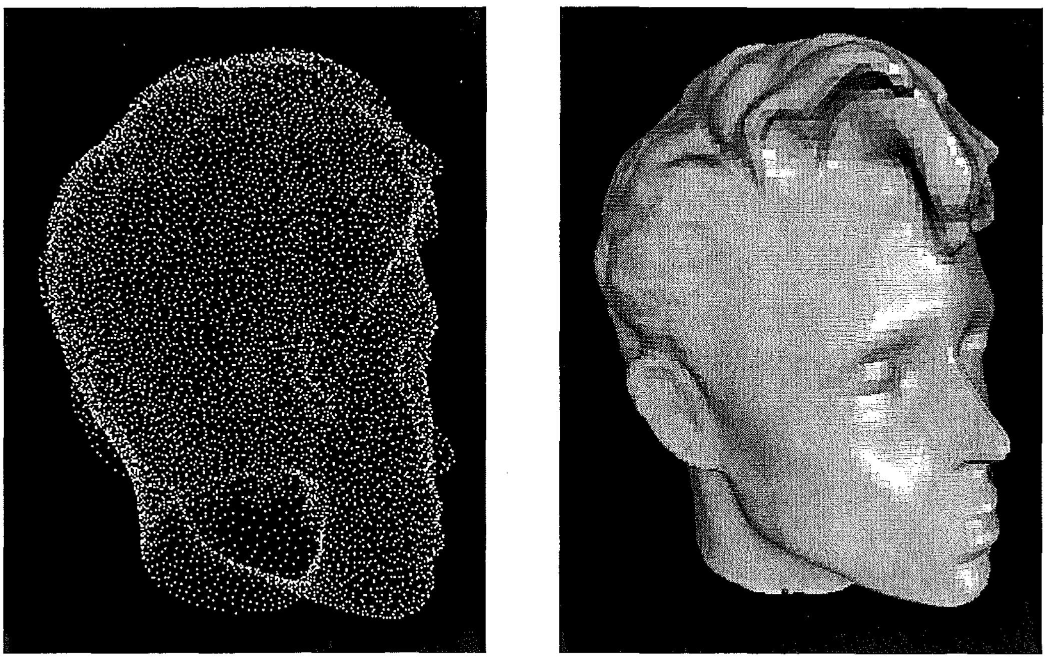
The problem considered in this chapter is

Problem 4.1. *Given n distinct points $\{(x_i, y_i, z_i)\}_{i=1}^n$ on a surface M in $\mathbb{R}^3$, find a surface M' that is a reasonable approximation to M.*

What is "reasonable"? This has been left deliberately vague as there are many possible ways in which the solution to this problem may be used and each may define its own meaning of "reasonable." Furthermore the set of data points $\{(x_i, y_i, z_i)\}_{i=1}^n$ controls how well M may be approximated. If all the data points come from the intersection of a sphere and the surface of a head, say, it may be perfectly valid to produce either surface as a solution.

Problem 4.1 is motivated by the need to process the data that comes from any of the many 3D range scanners that are currently available. They are capable of measuring the locations of tens or hundreds of thousands of points on the surface of an object. An example of one, the *Polhemus FastSCAN*, is shown in Figure 4.1. Problem 4.1 is neatly summarized in Figure 4.2:

Figure 4.1: A *Polhemus FastSCAN* laser scanner is used to measure the locations of thousands of points on the surface of an object. The computer is displaying a preview surface to show the operator what parts of the surface have been scanned. This preview surface can be used to estimate surface normals. (Picture courtesy *The Press.*)



(a) Input: A cloud of points. (b) Output: A fitted surface.

Figure 4.2: Graphical representation of Problem 4.1: From the cloud of points (Figure 4.2(a)) reproduce a resonable approximation to the orginal surface (Figure 4.2(b)). Images reproduced courtesy of Applied Research Associates NZ Ltd (see `http://www.aranz.co.nz/`).

Figure 4.2(a) is a typical example of the input to the problem under consideration and the required output corresponding to this particular input data set is shown in Figure 4.2(b).

The minimal output from these 3D scanners is a cloud of points. That is, a list of points in $\mathbb{R}^3$ each of which is on the surface of interest. However, as shown in Figure 4.1, the output from the *FastSCAN* displayed on screen is quite clearly more than just a cloud of points; there is a reasonable representation of a surface. This particular surface is just a preview surface and is generated by using information from the order in which points are scanned. However, this preview surface has many deficiencies. In particular it is usually in several distinct overlapping pieces and that each of these pieces may be self intersecting in a non-manifold way. So, while this preview surface is a reasonable approximation to the original surface for previewing the scanned data, it is not a reasonable approximation for many other applications. Despite this, there is still useful information to be obtained from this preview surface, in particular for each data point an estimate of the surface normal can be obtained. As shown below, these surface normals allow much better calculation of the surface.

Because not all scanners give this normal information, we look to solve this surface fitting problem with and without the normal information. As will be seen, these two cases make very little difference to the bulk of the algorithm.

4.1 Surface representation

As the solution to Problem 4.1 is to be used by a computer, *e.g.*, in an animation or with a CAD/CAM package, some consideration needs to be taken as to the form this surface should take. In computer graphics, there are two standard ways of representing a surface: implicit surfaces and parametric surfaces.

Implicit surfaces

An *implicit surface* is the zero set of a function of three variables, *i.e.*, it is the solution set to an equation like

$$f(x, y, z) = 0.$$

The function f that implicitly defines a surface is sometimes called an implicit function.

A closed surface divides $\mathbb{R}^3$ into three regions: the surface itself, a bounded region, called the inside or the interior, and an unbounded region called the outside or the exterior [23, XVII.2.4]. Thus a natural way to define a closed surface is in terms of its exterior and interior. A function that does this is sometimes called a *inside-outside* function and takes one sign inside the surface and the other sign outside the surface.

A particularly good choice for f is a *signed distance function.* If f is a signed distance function then the magnitude of $f(x, y, z)$ is the distance from the point (x, y, z) to the nearest point on the surface. The sign of $f(x, y, z)$ determines whether the point (x, y, z) is in the interior or the exterior of the surface. Note that the distance need not be Euclidean distance. The RBF interpolants that are computed below are approximations to a signed distance function.

While it is possible, and in fact quite common, to ray trace an implicit surface (to obtain a picture of the surface) [16, Chp. 5], we have not yet investigated this possibility for the large RBF functions used below to implicity define surfaces.

Parametric surfaces

For an implicit function, it is easy to determine if a given point is on the surface. However, in general it is difficult to generate points on the surface. With parametric surfaces this is reversed: generating points on the surface is easy but determining if a given point is on the surface is, in general, difficult.

Parametric surfaces in $\mathbb{R}^3$ are given by three bivariate functions. Given a domain $D \subseteq \mathbb{R}^2$, a parametric surface is

$$\left\{ \left(x(s,t), y(s,t), z(s,t) \right) : (s,t) \in D \right\}. \tag{4.1}$$

As it is difficult to parameterize a general surface, a parametric surface is normally made up of patches, with each patch given by an expression of the form (4.1).

A subset of parametric surfaces that are particular easy to render are the *faceted* surfaces. For a faceted surface each patch has a polygonal, often triangular or quadrilateral, boundary with linear or piecewise linear functions on each patch. When there are only triangular patches and linear functions then this is a *triangulation* of the surface.

Isosurfacing: conversion from implicit to parametric

An alternative to ray tracing for the display of implicit surfaces is to convert the surface into parametric form, particularly a faceted approximation to the surface. This is sometimes called *isosurfacing* since on the surface that is found the function that implicitly defines the surface is constant. The most popular technique for doing this conversion is probably

Marching Cubes [15, 44, 50]. Marching Tetrahedra [65] and Marching Triangles [36, 37, 38] are two other algorithms which perform this conversion.

Marching Cubes

Marching cubes was first introduced for isosurfacing medical data such as computed tomography (CT), magnetic resonance (MR), and single photon emission computed tomography (SPECT) [44]. This data is in the form of a grid of scalar values and Marching Cubes is thus based on gridded representation of $\mathbb{R}^3$. The surface intersects any voxel in the grid that has vertices that are differently signed. As there are only finitely many ways to triangulate a surface based on how it intersects a cube, Marching Cubes exploits a lookup table to triangulate the surface and thus produce a faceted approximation to the isosurface.

Marching cubes is easily applied to general implicit surfaces by evaluating the corresponding function on $\mathbb{R}^3$ at the points of an appropriate grid. A surface following algorithm can be used to ensure the function is only evaluated at the corners of voxels that intersect the surface. It should be clear that using a surface following technique is, in general, computationally more efficient than evaluating the function on the whole grid. However, if the function defines two or more separate surfaces then only one piece may be found by the surfacing following version while whole grid evaluation will find all surface pieces (to within the resolution of the grid).

A variant on Marching Cubes is Marching Tetrahedra [65]. As the name suggests, the principal difference between the two is that each of the voxels in the Marching Tetrahedra is a tetrahedron. A tetrahedron has nicer triangulation properties than a cube as there are only two ways to triangulate an intersecting surface rather than 14. However the useability of Marching Tetrahedra is limited to those functions that can be evaluated on the tetrahedral mesh. Since a cube may be decomposed into a tetrahedra, Marching Tetrahedra may be used on density data such as that from a CT scan. However, this implementation is essentially the same as Marching Cubes as the tetrahedral decomposistion of the cube just allows for a way to triangulate a surface intersecting a cube. Since RBFs may be evaluated at arbitary locations, Marching Tetrahedra is well suited to finding the isosurface of an RBF.

One pitfall of the Marching Cubes algorithm is the tendency it has to produce undesirable long skinny triangles. This is a result of how the voxels are triangulated. Marching Tetrahedra tends to produce a better surface in this respect than Marching Cubes does.

Marching Triangles

Marching Triangles [37] is a very different isosurfacing method from the two described above. It is not at all voxel based but instead attempts to successively add triangles to the boundary of the current faceted approximation. The *boundary* of a faceted surface is the set of all the edges that are incident on only one face.

Starting with an edge e, with vertices v_1 and v_2, Marching Triangles attempts to find a new vertex v_{new} such that the triangle $(v_1, v_2, v_{\text{new}})$ is a "good" addition to the surface. "Good" means that the triangle is close to equilateral and that the integrity of the surface is maintained, *e.g.,* each edge is incident on at most two faces and if any two faces intersect, it is only along an edge (these are the standard sort of conditions required by a triangulation of the plane).

How is the v_{new} found? First a projected vertex v_{proj} is chosen. This vertex is a fixed distance from the centre of e and in the same plane as the triangle incident on e. The new vertex, v_{new}, is then the closest point on the surface to v_{proj}. Before $(v_1, v_2, v_{\text{new}})$ is added to the surface, a check is made to ensure that it is a "good" triangle. If it is not, it is rejected.

If $(v_1, v_2, v_{\text{new}})$ is rejected, Marching Triangles tries to add a new triangle by considering the three vertices that belong to two consecutive edges on the boundary of the surface. Once again a check is made to ensure that any triangle that is added to the surface doesn't compromise the integrity of the surface.

Marching Triangles stops when there are no edges on the boundary (*i.e.,* the surface is closed) or some other user defined stopping criteria has been reached, *e.g.,* the boundary of the surface coincides with a user defined bounding box. Marching Triangles may also stop if it cannot add a triangle to any of the edges on the boundary.

4.2 Solution via RBFs

Our solution to Problem 4.1 is based on finding an RBF approximation to a signed distance function for the surface in question. Using a signed distance function to implicitly define a surface that interpolates a cloud of points is not a new idea. Hoppe *et al.* [40, 39] use a distance function based on tangent plane estimation while Bajaj *et al.* [4, 5] use a piecewise polynomial for this function. Muraki [52] attempts to minimize a norm that measures how well a "Blobby Model" (zero set of a sum of Gaussians) approximates the point cloud. This method does not necessarily produce a surface that interpolates all of the data points. Bernardini *et al.* [14] and

Bolle & Vemuri [17] review a number of techniques for solving the surface fitting Problem 4.1. Turk & O'Brien [67] use a technique which is essentially the same as ours but for surface design rather than reconstruction. Another difference from Turk & O'Brien is our emphasis on handling large data sets. Fast evaluation and fitting methods for RBFs are prerequisites if an RBF based approach is to handle data sets of the size generated by laser scanners. Turk & O'Brien also extend the idea of using a implicit description of a surface via an RBF to morphing one surface into another [66].

The following is an outline of our basic algorithm.

- Fit an RBF s to the data so that it in some sense represents the signed distance from the surface.
- Obtain a faceted approximation of the surface $\{x \in \mathbb{R}^d : s(x) = 0\}$ by using an isosurfacing technique

For the second of these two tasks, standard isosurfacing techniques such as marching cubes and marching triangles, discussed above, are sufficient. It is preferable that the signed distance function produced in step one have negative sign inside the surface. This makes the gradient of the RBF point in the same direction as the surface normals rather than being anti-parallel to them.

4.2.1 Implicit RBF interpolation

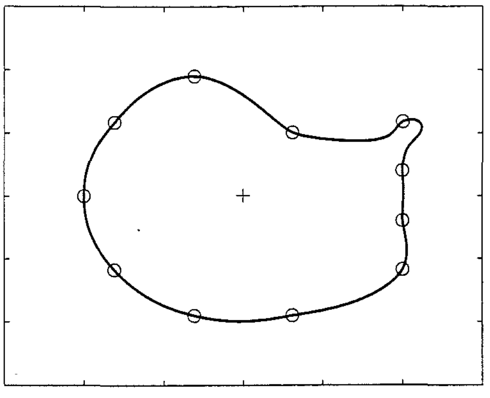
As mentioned above, we wish to find an RBF which implicitly defines the surface of interest. Because the corresponding linear system is invertible, solving the interpolation problem to the data

$$s(x_i, y_i, z_i) = 0, \qquad i = 1, \ldots, n,$$

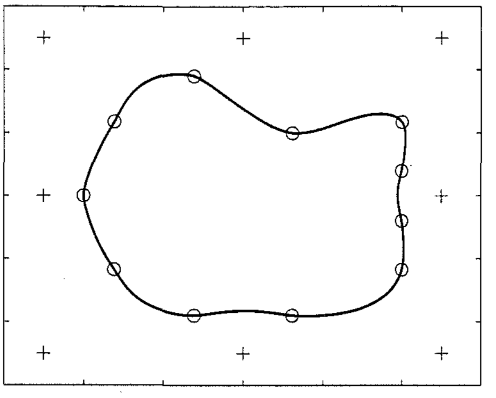
where $\{(x_i, y_i, z_i)\}_{i=1}^{n}$ are the points in the input cloud, leads to the useless solution $s = 0$. To avoid this problem, off-surface points are appended to the input data and are given non-zero values to interpolate to. This gives a more useful interpolation problem: Find s such that

$$\begin{aligned} &s(x_i, y_i, z_i) = 0, && i = 1, \ldots, n && \text{(points } \textit{on} \text{ the surface)},\\ &s(x_i, y_i, z_i) = d_i \neq 0, && i = n+1, \ldots, N && \text{(points } \textit{off} \text{ the surface)}. \end{aligned}$$

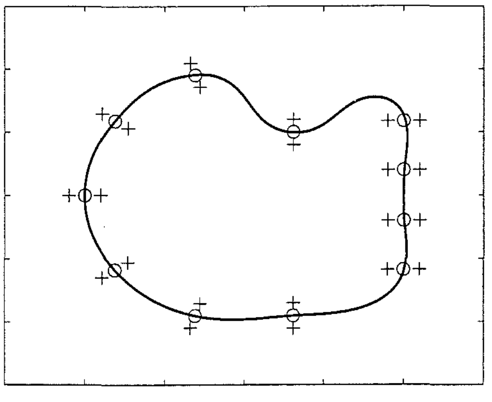
This still leaves the problem of determining the off-surface points $\{(x_i, y_i, z_i)\}_{i=n+1}^{N}$ and the corresponding interpolation values d_i.



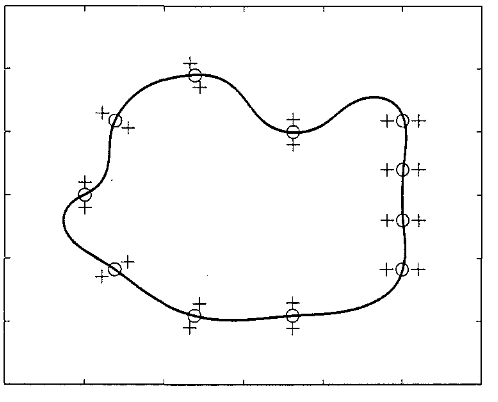
(a) Single interior point.



(b) Bounding box points.



(c) Using normal information.



(d) Effect of changing a normal.

Figure 4.3: Different strategies for off-surface or off-curve points. In the top two figures, both the on-curve points (∘) and the off-curve points (+) are used as centres for the RBF. In the lower two, only the off-curve points (+) are used as centres. In Figure 4.3(d), the normal corresponding to the left most point has been rotated through 90° from its position in Figure 4.3(c).

To demonstrate the various ways these off-surface points and interpolation values may be chosen, we lower the dimensions of the problem and look to fit a curve to a cloud of points in 2D. Three techniques for determining off-surface points are illustrated in Figure 4.3. In these examples the circles ($\circ$) designate the points in the original cloud and the crosses ($+$) designate the off-curve points that have been added. In each case it is just the off-curve points that have been changed.

Interior off-surface points

If points are known to be in the interior of the surface then they can be used as off-surface points with negative interpolation values. This is demonstrated in Figure 4.3(a) with a single off-surface interior point, ($+$). In general, for an arbitrary cloud of points we do not know where the interior of the surface is, making it difficult to find interior points. Choosing off-surface points in this way points is more appropriate to surface design rather than surface reconstruction [67].

Bounding box points

Exterior points are much easier to find than interior points. This is because the exterior of a surface is unbounded. If a simple surface such as a sphere or a rectangular box (cartesian product of intervals) encloses the cloud of points, then it is reasonable to expect that the exterior of the bounding surface is in the exterior of the surface defined by the cloud of points. Since a bounding box is easily calculated from the minimum and maximum coordinate values of the cloud of points, we normally use this bounding box to generate off-surface exterior points. Figure 4.3(b) illustrates this technique of our example 2D cloud of points. In this example eight points on the edges of a rectangle that bounds the cloud are used as off-curve points. The interpolation values d_i are determined by measuring the distance from the particular off-surface point to a closest point in the cloud, *i.e.*,

$$d_i = \min\{\mathrm{dist}(x_i, x_j) : \ j = 1, \ldots, n\}, \qquad i = n+1, \ldots, N,$$

where $\mathrm{dist}(\cdot, \cdot)$ is some appropriate measure of distance, usually Euclidean distance.

Exploiting surface normals

The last two figures, Figures 4.3(c) and 4.3(d), are examples of using normal information to determine the off-surface points. Since the surface normals point away from the surface,

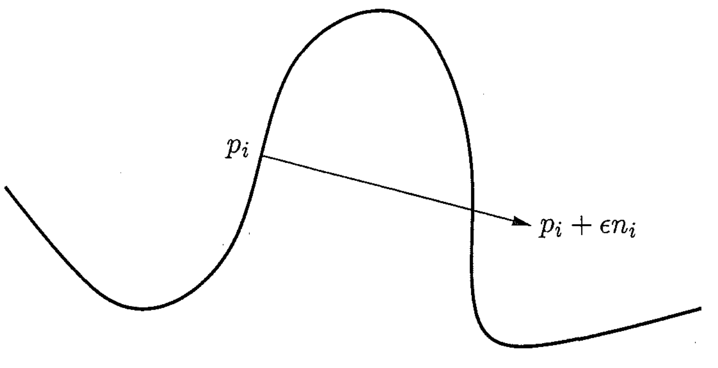


Figure 4.4: Projecting too far along a normal.

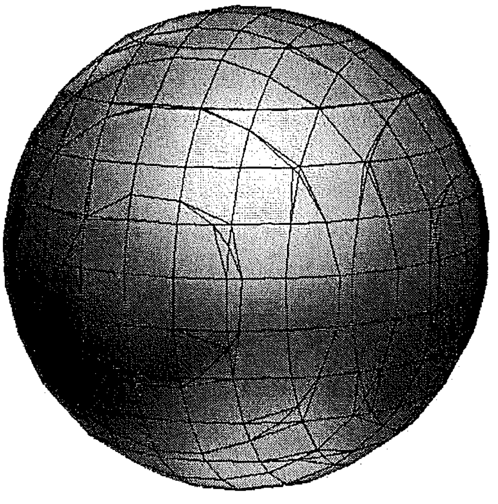
moving in this direction from the data point leads to an off-surface point. In other words, if n_i is the surface normal at a data point $p_i = (x_i, y_i, z_i)$, then $p + \epsilon n_i$ can be used as an off-surface point. As the interpolant is to approximate a signed distance function, the corresponding interpolation value d_i is chosen as a multiple of ϵ. The value of ϵ is normally left as a user defined parameter. It must be large enough to overcome any noise in the data and small enough not to cross to the other side of the surface in concave regions, as illustrated in Figure 4.4.

Experience has shown that it is better to replace a data point p_i with two off-surface points, one outside the surface $p + \epsilon n_i$ and one inside $p - \epsilon n_i$. The interpolation values are then chosen to reflect whether the off-surface points are interior or exterior: $d_i = \epsilon$ for exterior and $d_i = -\epsilon$ for interior. This has been done in Figures 4.3(c) and 4.3(d). Note that no original data points ($\circ$) were used in the interpolation problem. Figure 4.3(d) shows the effect on the curve of rotating one of the estimated surface normals.

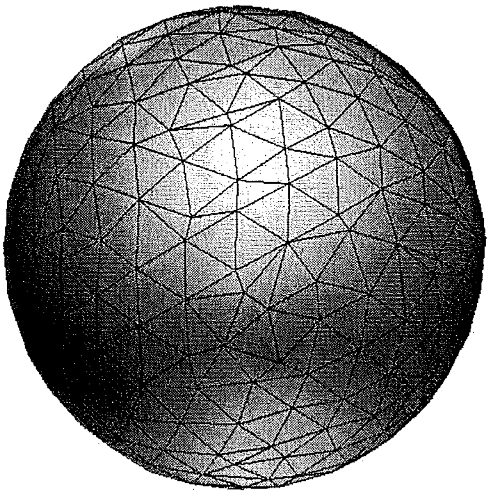
4.3 Examples

In this section we present some examples of how well this technique works on various data sets, Table 4.1. These examples are divided into two categories: synthetic surfaces and real world surfaces. The synthetic surfaces are various mathematically generated surfaces and are designed to show specific aspects of the technique. The real world examples used range scanner data from actual objects.

All data sets were fitted using biharmonic radial basis functions *i.e.*, $\phi(r) = r$ was used as the basic function. The fitting was done using an extension of the Domain Decomposition method of Beatson, Light & Billings [11]. Three isosurfacing programs were used to produce



(a) Marching Cubes output.



(b) Marching Triangles output

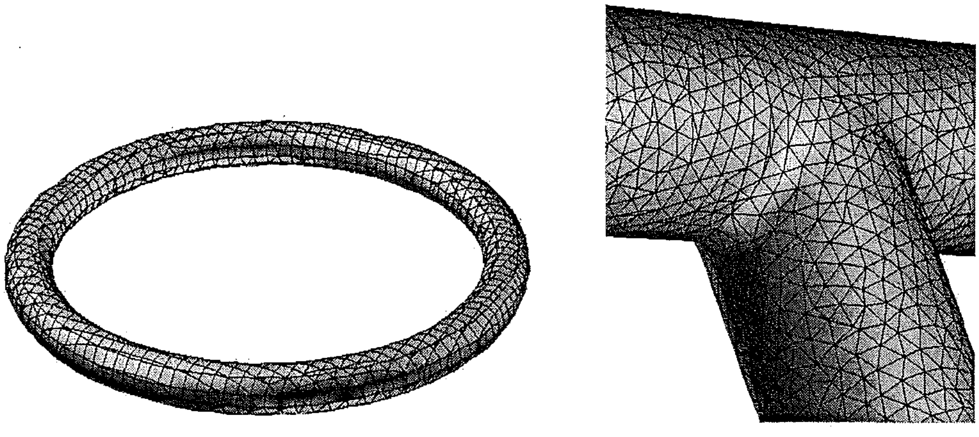
Figure 4.5: Marching Cubes versus Marching Triangles

the faceted surfaces in this chapter

- *MCubes* is an implementation of Marching Cubes. MCubes uses normal information to ensure the correct local convexity of the surface is obtained. One consequence of this is that as finer grids are taken, the surface area of the faceted surface tends to the true area of surface.
- *Saunter* is an implementation of Marching Triangles.
- *FastSurf* is an implementation of Marching Tetrahedra.

4.3.1 Synthetic surfaces

The first example is the sphere in Figure 4.5. In this example, as with the next couple of examples, a bounding box has been used to generate the off-surface points. This technique has no difficulty with simple shapes such as this. As with most of the examples in this chapter, the surfaces are displayed with the triangles and quadrilaterals which make up the faceted surface. This example clearly shows the difference between Marching Cubes and Marching Triangles. The sphere in Figure 4.5(a) has been faceted by Marching Cubes and the long skinny triangles typical of this method are quite evident. The triangles in Figure 4.5(b), produced by Marching Triangles, are, by contrast, more regular in shape.



(a) Torus

(b) A nice join between two cylinders

Figure 4.6: Aspects of implicit surfaces of RBFs.

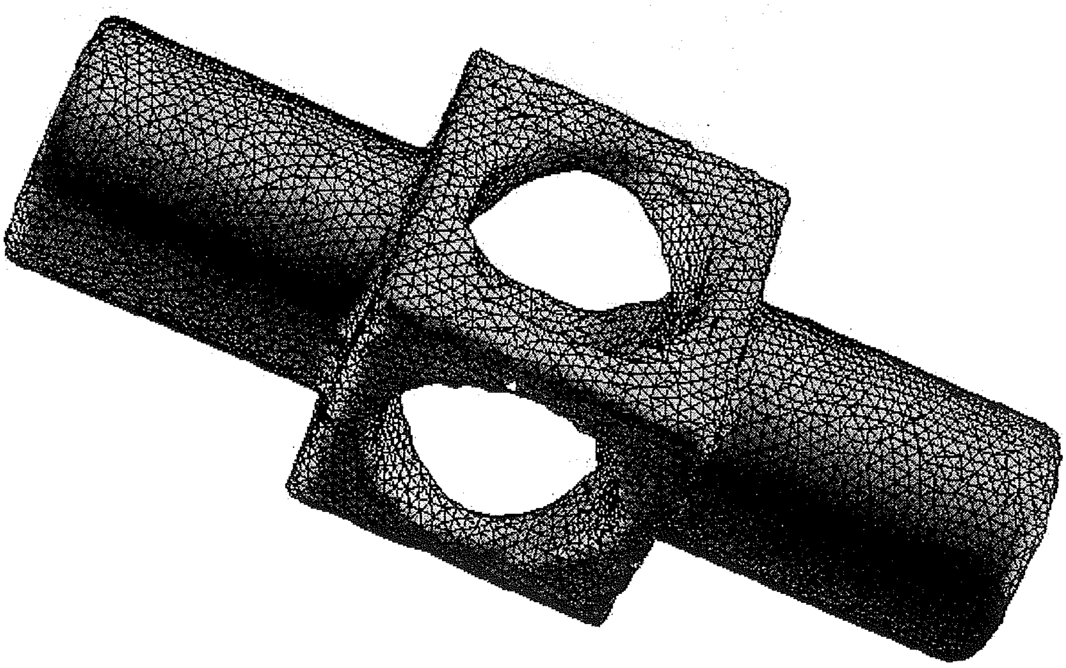
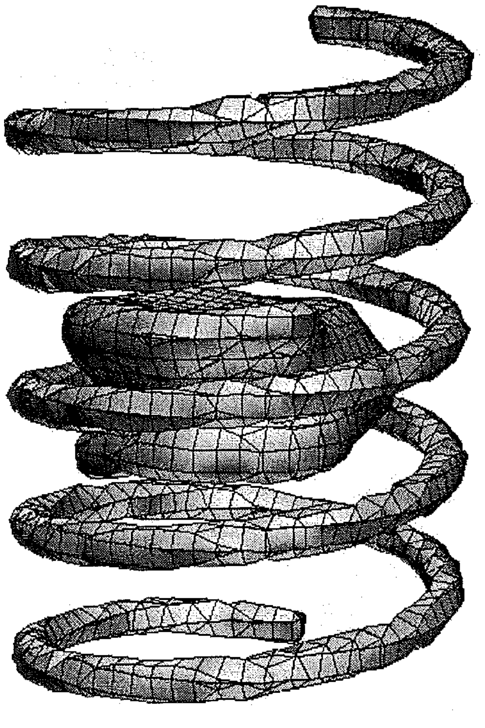
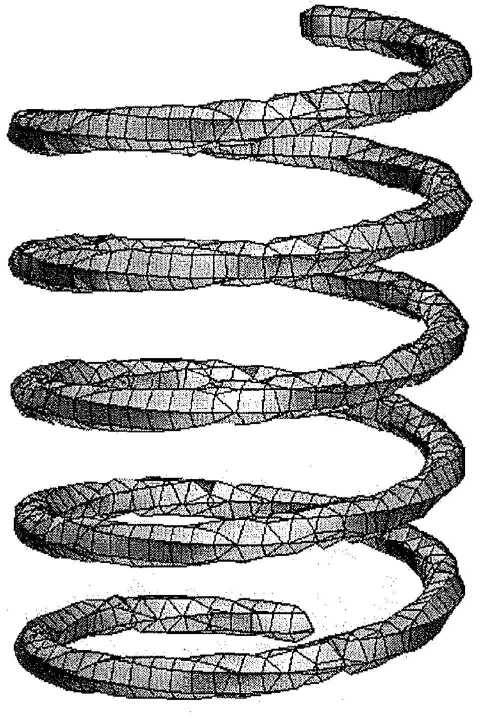


Figure 4.7: A CSG object made up of two cylinders and a cube with a sphere removed from its centre.



(a) Using a bounding box.



(b) Using surface normals.

Figure 4.8: Bounding box versus normals

More complicated surfaces are shown in Figure 4.6. The torus in Figure 4.6(a) shows the ability of our technique to handle objects with holes. Figure 4.6(b) shows a surface produced by two cylinders intersecting at right angles. There is clearly a smooth transition from one cylinder to the other. As evidenced by Figure 4.7, complicated surfaces are handled with ease, showing no difficulty in the transition between flat and curved parts of the surface.

While the bounding box technique works quite well for releativily simple objects, it fails on more complicated objects. Figure 4.8(a) shows a typical situation where bounding box points fail to produce a satisfactory surface. The blob in the centre of the helix is an artefact of the method; all of the data points lie on the surface of the helix. This sort of artefact tends to occur when there are large regions exterior to the surface but inside the cloud of data points and hence the bounding box. Figure 4.8(b) shows the corresponding surface produced using normals to generate off surface points.

4.3.2 Real world surfaces

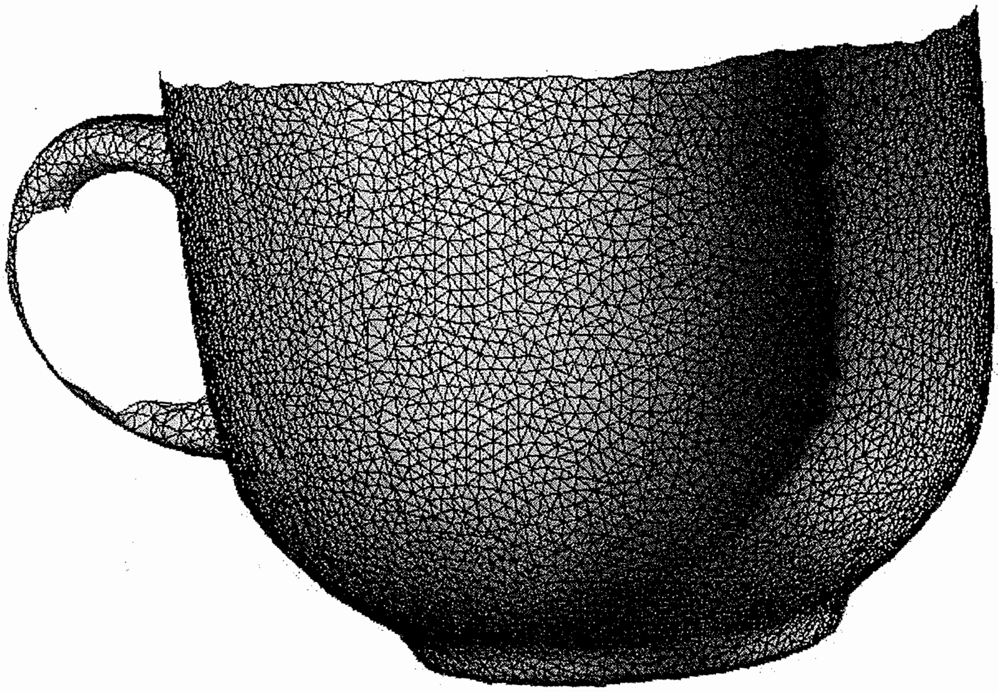
This subsection presents a number of surfaces generated from range scanning of real world objects. In the first two examples, Figure 4.9 and 4.10, the scan information is incomplete. Data is missing from the inside of the fingers in Figure 4.9(a) and from parts of the handle in Figure 4.10(a). These two figures show the preview surfaces produced by the *FastSCAN* after some post-processing. While the fingers in Figure 4.9(b) have been reproduced quite well despite the absence of data, the handle in Figure 4.10(b) has a protrusion that is not part of the original surface. Both of these examples were fitted using bounding box off-surface points.

The remaining examples, Figures 4.11–4.17 show that the technique works well for a variety of surfaces, both biological and artificial. Note that sharp edges tend to be rounded off, especially in the distributor cap, Figure 4.12 and the engine part Figure 4.15.

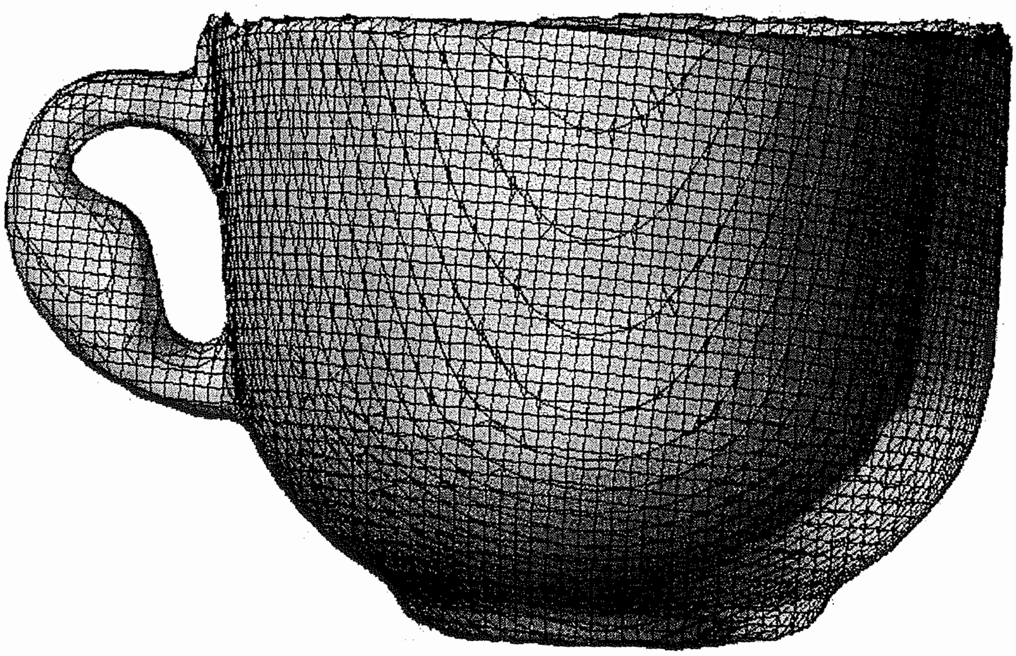
The final example of a dragon, Figure 4.17, shows the result of applying the method to a very large data set of 437,645 points. To process this data every third point was replaced by a pair of normal based off-surface points, *i.e.,* there were 568,938 nodes in the interpolation problem. An RBF was fitted to this data using a reduction technique resulting in an RBF with 72,461 centres. Fitting and evaluating such a function is computationally very expensive. It is not possible on any but the most well-endowed computers without using the fast methods that have been described in this thesis.

(a) Original scan of a hand. (b) Fitted surface.

Figure 4.9: Filling in gaps: Hand.

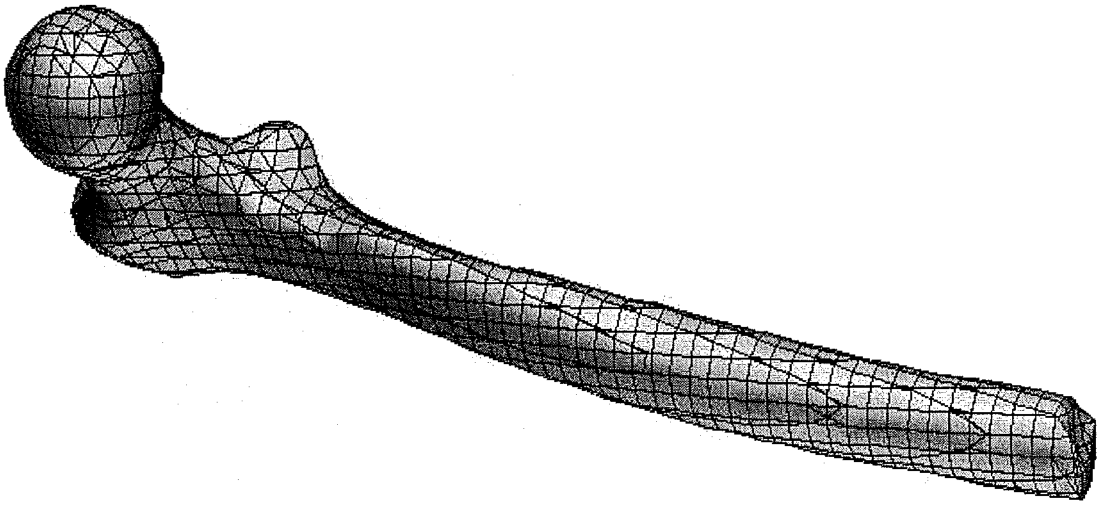


(a) Original scan of a cup.

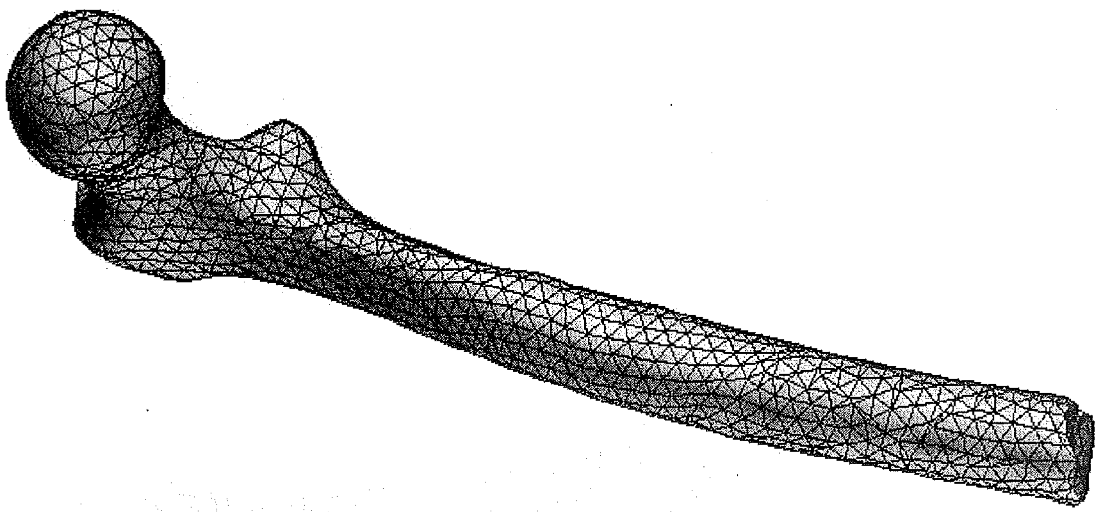


(b) Fitted surface.

Figure 4.10: Filling in gaps: Cup.



(a) Marching Cubes output



(b) Marching Triangles output.

Figure 4.11: Femur

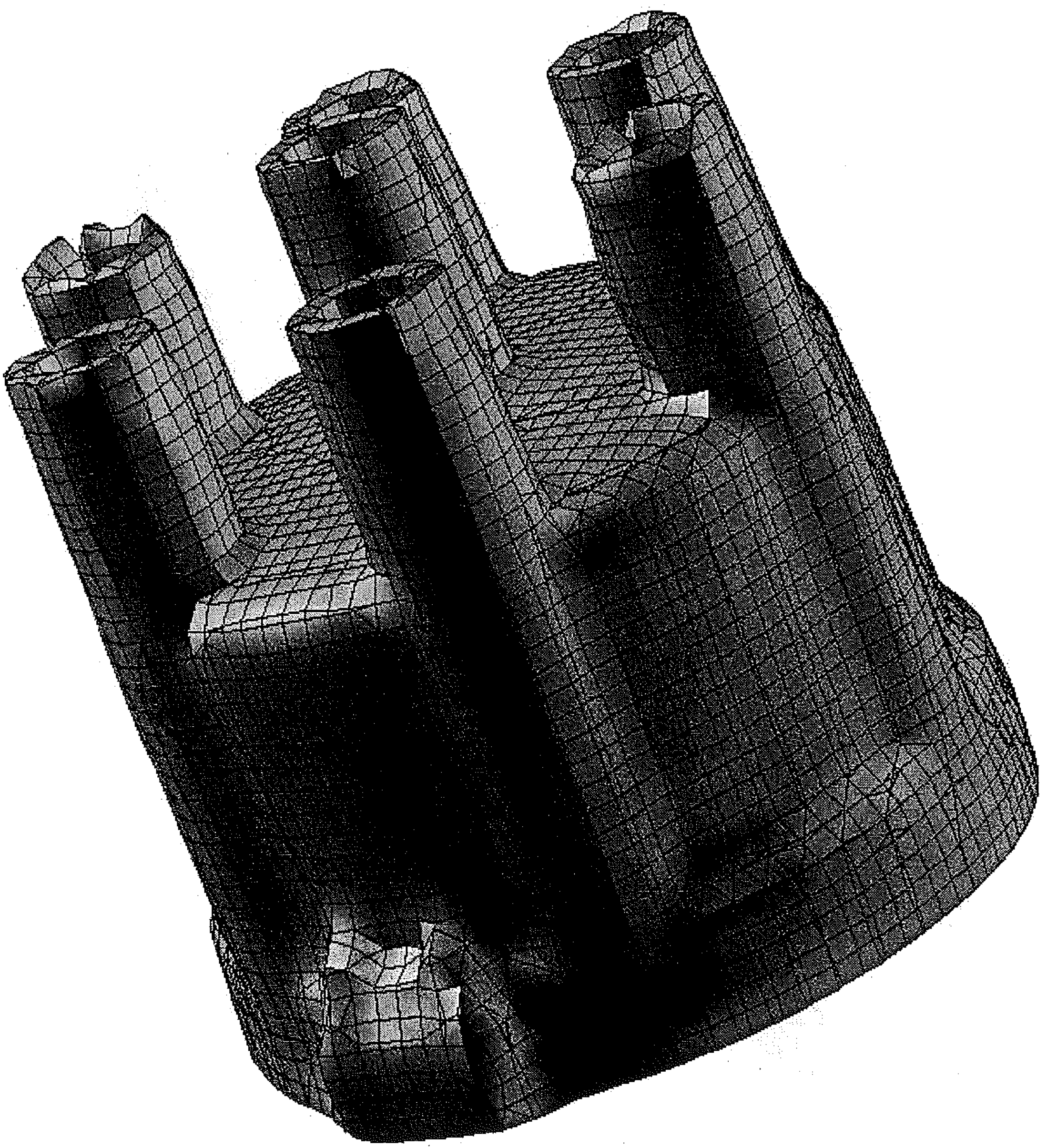


Figure 4.12: Distributor cap

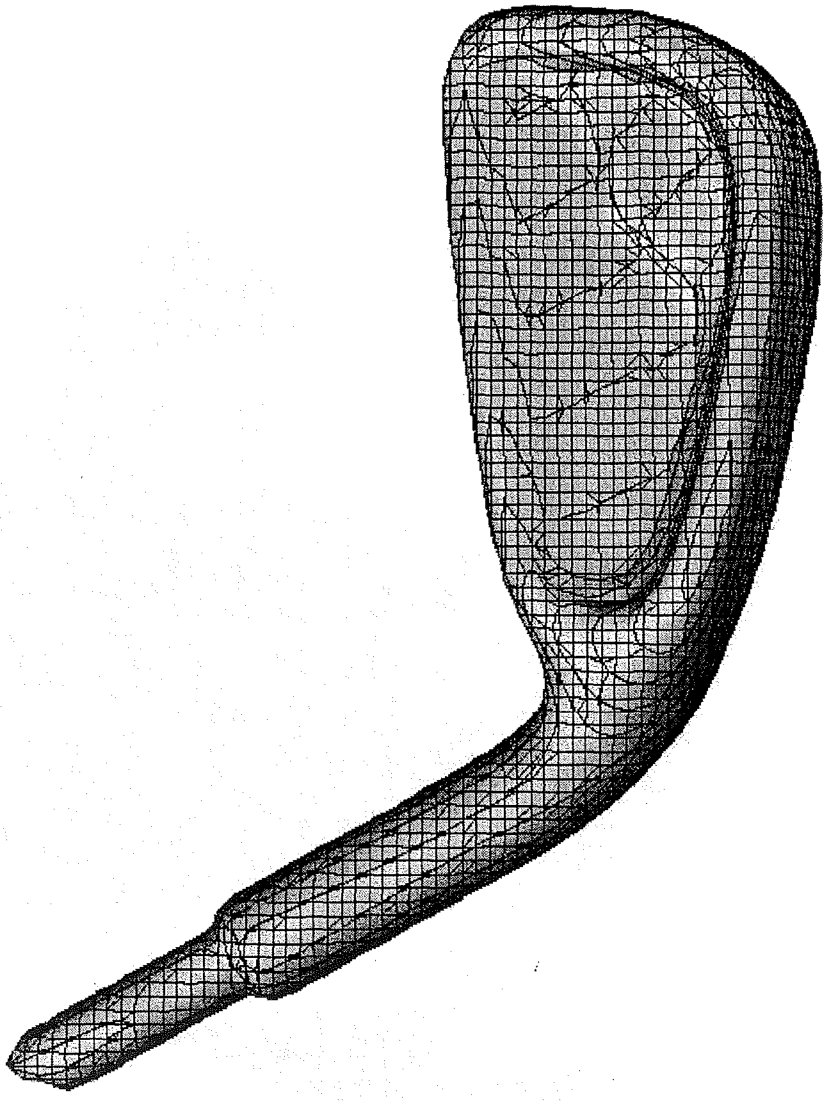


Figure 4.13: Golf club

	Figure	Original Data	Isosurfacer[a]	Final Surface		Source[b]
				Points	Facets	
Hand	4.9	13,340	MCubes	2,456	3,035	ARA
Cup	4.10	7,193	MCubes	15,550	19,234	ARA
Femur	4.11(a)	2,385	MCubes	1,848	2,266	ARA
	4.11(b)	2,385	Saunter	1,563	3,121	ARA
Distributor cap	4.12	12,745	MCubes	6,667	7,370	UWash
Golf club	4.13	16,864	MCubes	6,022	7,303	UWash
Teapot	4.14	26,103	Saunter	8,753	17,497	UWash
Engine part	4.15	30,937	MCubes	22,612	24,709	UWash
Scapula	4.16	25,183	FastSurf	4,928	9,852	ARA
Dragon	4.17	437,645	FastSurf			Stanford

[a]MCubes = Marching Cubes, Saunter = Marching Triangles, FastSurf = Marching Tetrahedra.

[b]ARA = Applied Research Associates NZ Ltd., `http://www.aranz.co.nz/`,
UWash = Computer Graphics Group, University of Washington,
`http://www.cs.washington.edu/research/graphics/`,
Stanford = Stanford University Computer Graphics Laboratory,
`http://www-graphics.stanford.edu/data/3Dscanrep/`.

Table 4.1: Summary of scanned data

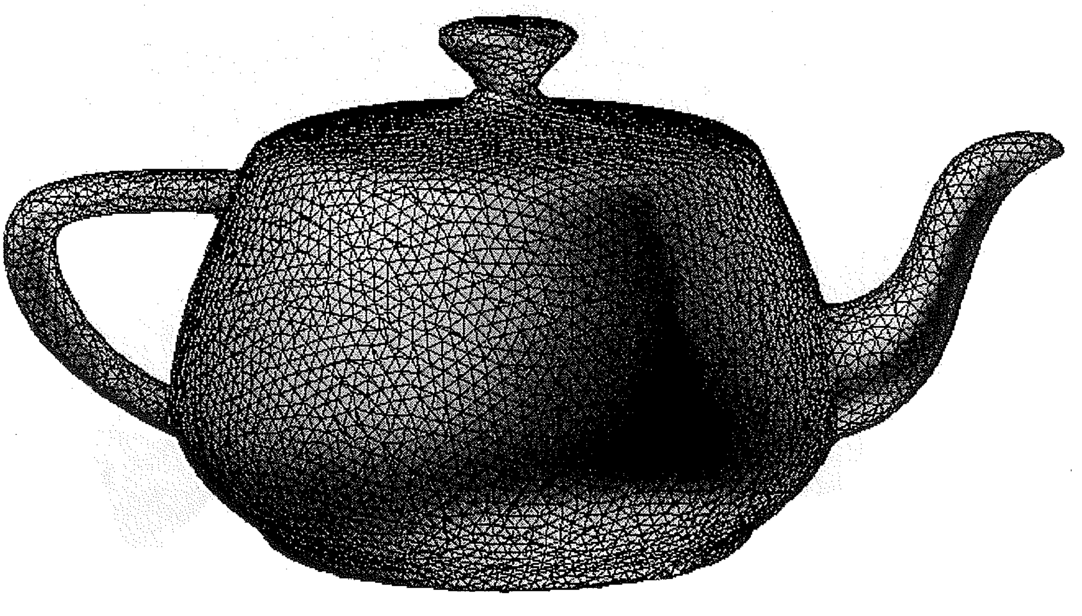


Figure 4.14: Teapot

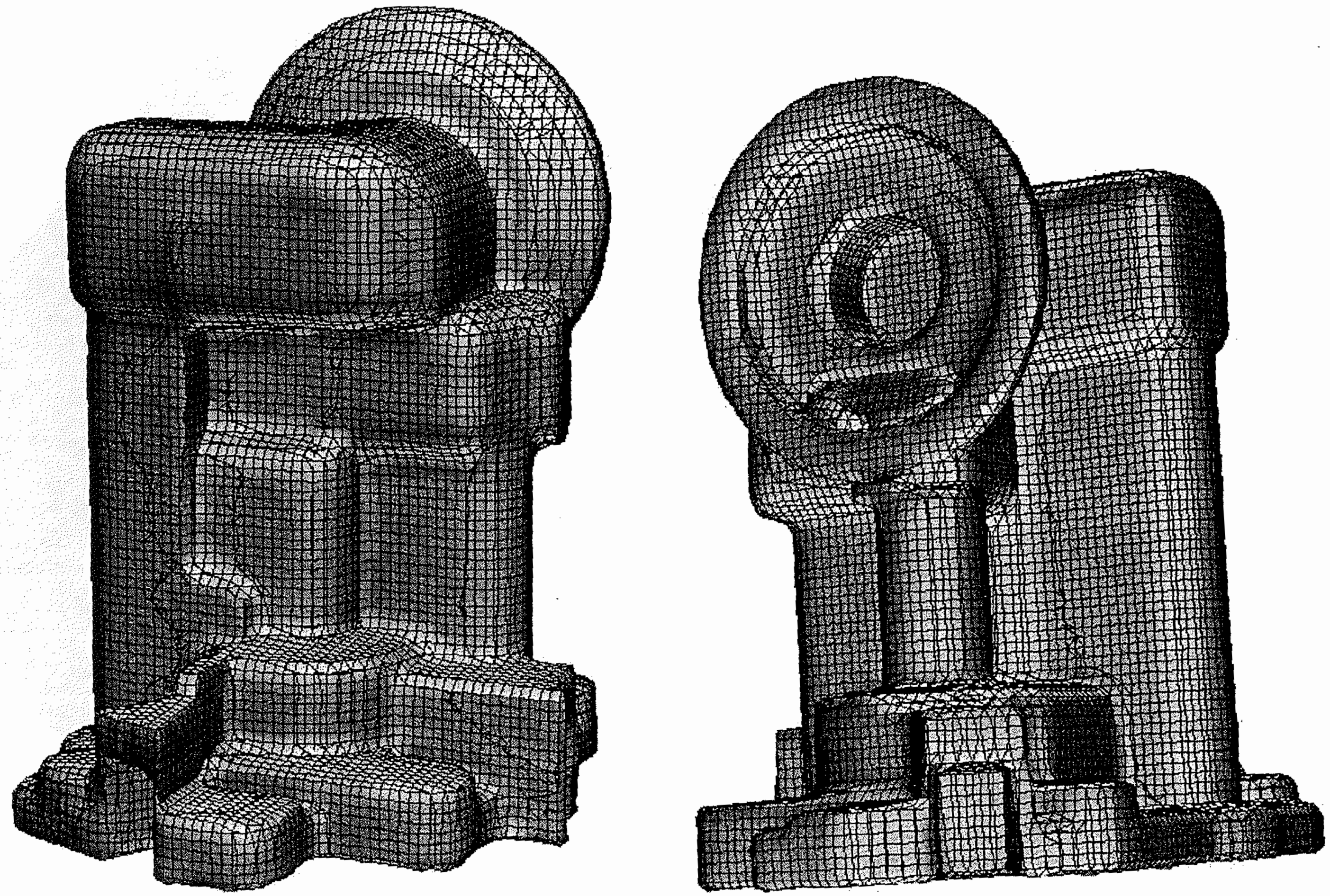


Figure 4.15: Engine part

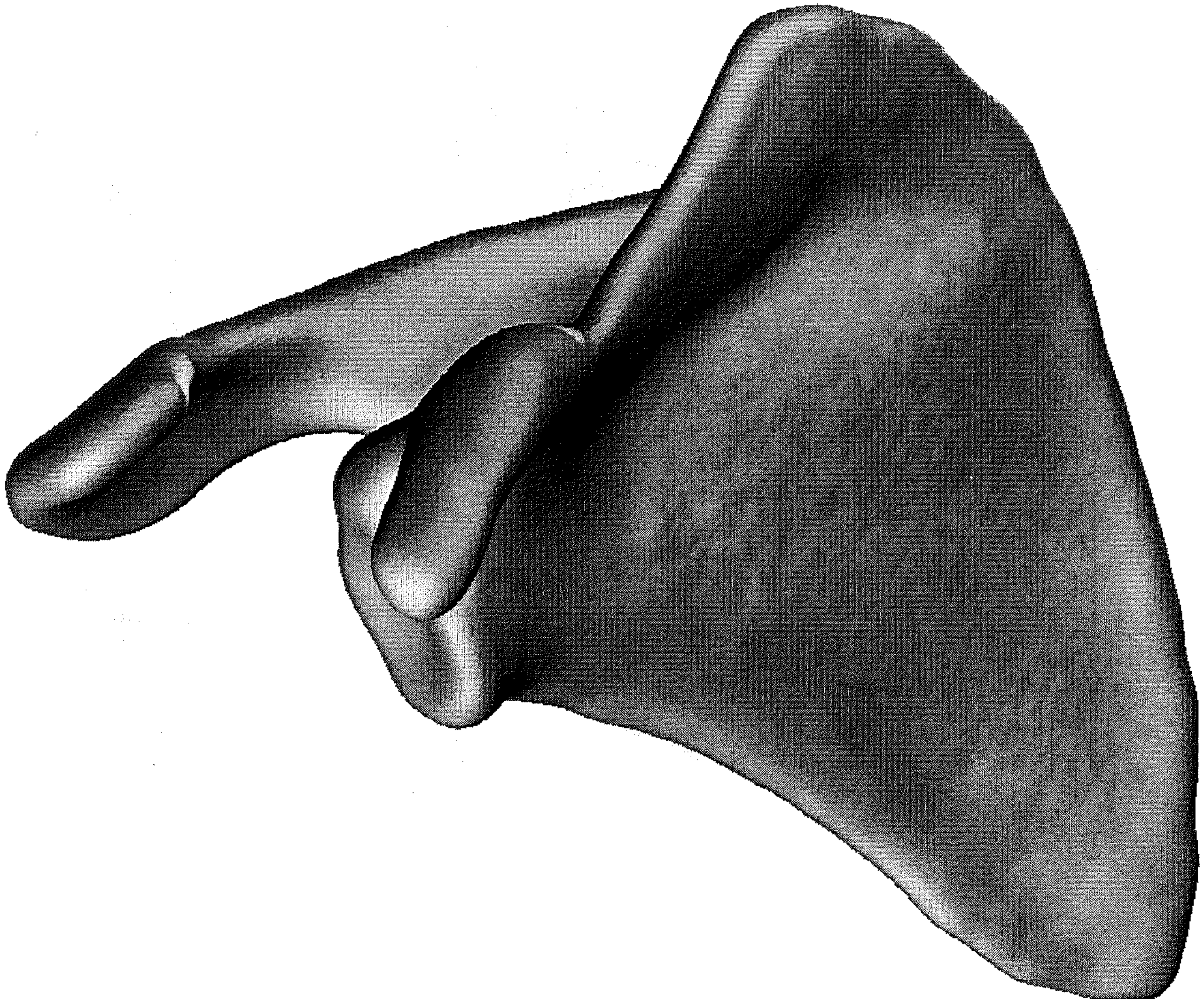


Figure 4.16: Scapula (shoulder blade)

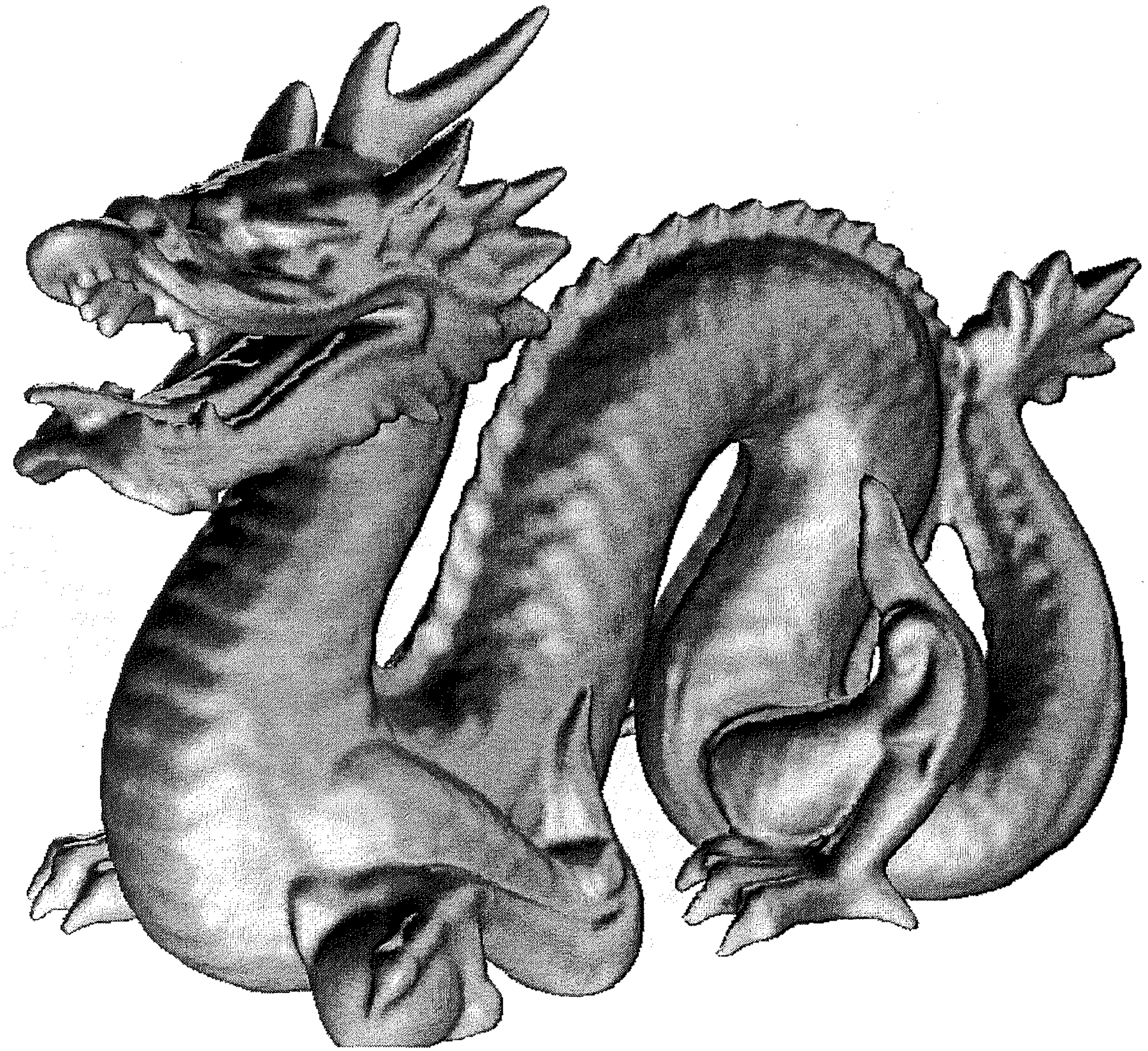


Figure 4.17: Dragon

Chapter 5

Summary and further work

Radial Basis Functions are well suited to solving interpolation problems to scattered data. However, for real world applications, data sets can be expected to be very large, *e.g.*, the dragon example in Chapter 4 has in excess of 70,000 data points. With one term for every data point, the RBF solutions to these problems are correspondingly large, making direct fitting and evaluation of such RBFs computationally intensive. However, the Fast Multipole Method may be applied to reduce this cost. Before the FMM may be applied, a range of analytical results are required for the corresponding basic function. These results have been developed for polyharmonic RBFs in 4 dimensions (Chp. 2) and multiquadrics in arbitrary dimensions (Chp. 3). It is possible that tighter error bounds may be achieved for the multiquadric by using a basis which is more appropriate than the monomial basis for the expansions, *e.g.*, the inner functions used for the polyharmonic expansions.

The technique discussed in Chapter 4 for fitting surfaces to point clouds has been shown by example to work well, even on large data sets. Biharmonic RBFs were used in these examples because they have proven to be good interpolants for other applications. The "niceness" of biharmonic interpolation is due in part to their characterization as minimum energy interpolants. However, the semi-inner product used in this characterization is probably less than optimal and even better results may be possible with a more appropriate norm.

Bibliography

[1] M. Abramowitz and I. A. Stegun, eds., *Handbook of Mathematical Functions*, Dover, New York, 1965.

[2] J. F. Adams, *Lectures on Lie Groups*, W. A. Benjamin, Inc., New York, 1969.

[3] S. L. Altmann, *Rotations, Quaternions and Double Groups*, Oxford University Press, 1986.

[4] C. L. Bajaj, F. Bernardini, and G. Xu, *Automatic reconstruction of surfaces and scalar fields from 3D scans*, in SIGGRAPH'95, 1995, pp. 109–118.

[5] ——, *Reconstructing surfaces and functions on surfaces from unorganized three-dimensional data*, Algorithmica, 19 (1997), pp. 243–261.

[6] R. K. Beatson, J. C. Carr, J. B. Cherrie, T. Evans, W. R. Fright, B. C. McCallum, and T. Mitchell, *Practical surface reconstruction with implicit functions.* In preparation, 2000.

[7] R. K. Beatson, J. B. Cherrie, and C. T. Mouat, *Fast fitting of radial basis functions: Methods based on preconditioned GMRES iteration*, Advances in Computational Mathematics, 11 (1999), pp. 253–270.

[8] R. K. Beatson, J. B. Cherrie, and D. L. Ragozin, *Fast evaluation of radial basis functions: Methods for four-dimensional polyharmonic splines*, SIAM J. Math. Anal., (2000). To appear.

[9] R. K. Beatson and L. Greengard, *A short course on fast multipole methods*, in Wavelets, Multilevel Methods and Elliptic PDEs, M. Ainsworth, J. Levesley, W. Light, and M. Marletta, eds., Oxford University Press, 1997, pp. 1–37.

[10] R. K. Beatson and W. A. Light, *Fast evaluation of radial basis functions: Methods for two-dimensional polyharmonic splines*, IMA Journal of Numerical Analysis, 17 (1997), pp. 343–372.

[11] R. K. Beatson, W. A. Light, and S. Billings, *Fast solution of the radial basis function interpolation equations: Domain decomposition methods*. Preprint, 1999.

[12] R. K. Beatson and G. N. Newsam, *Fast evaluation of radial basis functions: I*, Computers Math. Applic., 24 (1992), pp. 7–19.

[13] R. K. Beatson, A. M. Tan, and M. J. D. Powell, *Fast evaluation of radial basis functions: Methods for 3-dimensional polyharmonic splines*. In preparation.

[14] F. Bernardini, C. Bajaj, J. Chen, and D. R. Schikore, *Automatic reconstruction of 3D CAD models from digital scans*, International Journal on Computational Geometry and Applications, 9 (1999), p. 327.

[15] J. Bloomenthal, *Polygonization of implicit surfaces*, Computer Aided Geometric Design, 5 (1988), pp. 341–355.

[16] ——, ed., *Introduction to Implicit Surfaces*, Morgan Kaufmann, San Francisco, California, 1997.

[17] R. M. Bolle and B. C. Vemuri, *On three-dimensional surface reconstruction methods*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 13 (1991), pp. 1–13.

[18] J. C. Carr, W. R. Fright, and R. K. Beatson, *Surface interpolation with radial basis functions for medical imaging*, IEEE Trans. Medical Imaging, 16 (1997), pp. 96–107.

[19] E. W. Cheney and W. A. Light, *A Course in Approximation Theory*, Brooks Cole, Pacific Grove, 1999.

[20] J. B. Cherrie, R. K. Beatson, and G. N. Newsam, *Fast evaluation of radial basis functions: Methods for generalised multiquadrics in* $\mathbb{R}^n$, Tech. Rep. UCDMS2000/1, Dept. of Mathematics and Statistics, University of Canterbury, Private Bag 4800, Christchurch, New Zealand, January 2000.

[21] R. V. Churchill and J. W. Brown, *Complex Variables and Applications*, McGraw Hill, New York, 4th ed., 1984.

[22] J. DUCHON, *Splines minimizing rotation-invariant semi-norms in Sobolev spaces*, in Constructive Theory of Functions of Several Variables, W. Schempp and K. Zeller, eds., no. 571 in Lecture Notes in Mathematics, Berlin, 1977, Springer-Verlag, pp. 85–100.

[23] J. DUGUNDJI, *Topology*, Allyn and Bacon, Inc., Boston, 1966.

[24] N. DYN, D. LEVIN, AND S. RIPPA, *Numerical procedures for surface fitting of scattered data by radial functions*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 639–659.

[25] M. A. EPTON AND B. DEMBART, *Multipole translation theory for the three-dimensional Laplace and Helmholtz equations*, SIAM J. Sci. Comput., 16 (1995), pp. 865–897.

[26] J. FLUSSER, *An adaptive method for image registration*, Pattern Recognition, 25 (1992), pp. 45–54.

[27] R. FRANKE, *Scattered data interpolation: Tests of some methods*, Mathematics of Computation, 38 (1982), pp. 181–200.

[28] M. GOLOMB AND H. F. WEINBERGER, *Optimal approximation and error bounds*, in On Numerical Approximation, R. E. Langer, ed., Madison, 1959, Unversity of Wisconsin Press, pp. 117–190.

[29] L. GREENGARD AND V. ROKHLIN, *A fast algorithm for particle simulations*, J. Comput. Phys, 73 (1987), pp. 325–348.

[30] K. GUO, S. HU, AND X. SUN, *Conditionally positive definite functions and Laplace-Stieltjes integrals*, J. Approx. Theory, 74 (1993), pp. 249–265.

[31] R. L. HARDER AND R. N. DESMARAIS, *Interpolation using surface splines*, J. Aircraft, 9 (1972), pp. 189–191.

[32] R. L. HARDY, *Multiquadric equations of topography and other irregular surfaces*, J. Geophysical Research, 76 (1971), pp. 1905–1915.

[33] ——, *Theory and applications of the multiquadric biharmonic method*, Comput. Math. Applic., 19 (1990), pp. 163–208.

[34] G. HEINDL, *Interpolation and approximation by piecewise quadratic C^1 functions of two variables*, in Multivariate Approximation Theory, W. Schempp and K. Zeller, eds., Birkhäuser Verlag, 1979, pp. 146–161.

[35] E. Hewitt and K. A. Ross, *Abstract Harmonic Analysis, II*, vol. 152 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, Berlin, 1970. QA 403 .H611 vol. 2.

[36] A. Hilton, A. J. Stoddart, J. Illingworth, and T. Windeatt, *Building 3D models of complex objects*, in 14th Eurographics Conference, H. Jones, R. Raby, and D. Vicars, eds., vol. 1, London, 1996, pp. 193–204.

[37] ——, *Marching triangles: Range image fusion for complex object modelling*, in IEEE International Conference on Image Processing, P. Delogne, ed., vol. II, Lausanne, Switzerland, 1996, pp. 381–384.

[38] ——, *Reliable surface reconstruction from multiple range images*, in Fourth European Conference on Computer Vision, B. Buxton and R. Cipolla, eds., Cambridge, 1996, pp. 117–126.

[39] H. Hoppe, T. DeRose, T. Duchamp, M. Halstead, H. Jin, J. McDonald, J. Schweitzer, and W. Stuetzle, *Piecewise smooth surface reconstuction*, in Proceedings of the 21st Annual Conference on Computer Graphics, 24–29 July, 1994, Orlando, Florida, 1994, pp. 295–302.

[40] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle, *Surface reconstruction from unorganized points*, Computer Graphics (SIGGRAPH'92 proceedings), 26 (1992), pp. 71–78.

[41] A. Kyrala., *Theoretical Physics: Applications of Vectors, Matrices, Tensors and Quaternions*, W. B. Saunders Company, Philadelphia, 1967.

[42] J. Levesley, W. Light, D. Ragozin, and X. Sun, *A simple approach to the variational theory for interpolation on spheres*, in New Developments in Approximation Theory, M. W. Müller, M. D. Buhmann, D. H. Mache, and M. Falten, eds., Birkhäuser Verlag, 1999, pp. 117–143.

[43] W. Light, *Variational methods for interpolation, particularly by radial basis functions*, in Numerical Analysis 1995, D. F. Griffiths and G. A. Watson, eds., London, 1996, Longam, pp. 94–106.

[44] W. E. Lorensen and H. E. Cline, *Marching cubes: A high resolution 3D surface construction algorithm*, Computer Graphics, 21 (1987), pp. 163–169.

[45] J. C. MAIRHUBER, *On Haar's theorem concerning Chebyčhev approximation problems having unique solutions*, Proc. Amer. Math. Soc., 7 (1956), pp. 609–615.

[46] S. MCCULLOUGH AND D. WULBERT, *The topological spaces that support Haar systems*, Proc. Amer. Math. Soc., 94 (1985), pp. 687–692.

[47] J. MEINGUET, *An intrinsic approach to multivariate spline interpolation at arbitary points*, in Polynomial and Spline Approximation, B. N. Sahney, ed., NATO Advanced Study Institute Series, Dordrecht, 1979, D. Reidel Publishing, pp. 163–190.

[48] ——, *Multivariate interpolation at arbitary points made simple*, Zeitschrift fur Angewante Mathematik and Physik, 30 (1979), pp. 292–304.

[49] C. A. MICCHELLI, *Interpolation of scattered data: Distance matrices and conditionally positive definite functions*, Constr. Approx., 2 (1986), pp. 11–22.

[50] C. MONTANI, R. SCATENI, AND R. SCOPIGNO, *Decreasing isosurface complexity via discrete fitting*, Computer Aided Geometric Design, 17 (2000), pp. 207–232.

[51] C. MÜLLER, *Spherical Harmonics*, vol. 17 of Lecture Notes in Mathematics, Springer-Verlag, New York, 1966.

[52] S. MURAKI, *Volumetric shape description of range data using "blobby model"*, Computer Graphics, 25 (1991), pp. 227–235.

[53] M. A. NAIMARK AND A. I. STERN, *Theory of Group Representations*, vol. 246 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, New York, 1982.

[54] M. J. D. POWELL, *The theory of radial basis function approximation in 1990*, in Advances in Numerical Analysis II: Wavelets, Subdivision Algorithms and Radial Functions, W. Light, ed., Oxford University Press, 1992, pp. 105–210.

[55] D. L. RAGOZIN, *Constructive polynomial approximation on spheres and projective spaces*, Trans. Amer. Math. Soc., 162 (1971), pp. 157–170.

[56] ——, *Uniform convergence of spherical harmonic expansions*, Math. Ann., 195 (1972), pp. 87–94.

[57] Y. SAAD AND M. H. SCHULTZ, *GMRES: A generalized minimal residual algorithm for solving nonsymetric linear equations*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.

[58] R. SCHABACK, *Comparison of radial basis function interpolants*, in Multivariate Approximations: From CAGD to Wavelets, K. Jetter and F. Utreas, eds., Singapore, 1993, World Scientific, pp. 293–305.

[59] ——, *Multivariate interpolation and approximation by translates of a basis function*, in Approximation Theory VIII, Vol. 1: Approximation and Interpolation, C. K. Chui and L. L. Schumaker, eds., World Scientific, 1995, pp. 491–514.

[60] L. L. SCHUMAKER, *On the dimension of spaces of piecewise polynomials in two variables*, in Multivariate Approximation Theory, W. Schempp and K. Zeller, eds., Birkhäuser Verlag, 1979, pp. 396–412.

[61] L. SCHWARTZ, *Théorie des Distributions*, Hermann, Paris, 1966.

[62] R. SIBSON AND G. STONE, *Computation of thin-plate splines*, SIAM Journal on Scientific and Statistical Computing, 12 (1991), pp. 1304–1313.

[63] J. STEWART, *Positive definite functions and generalizations, an historical survey*, Rocky Mountain Journal of Mathematics, 6 (1976), pp. 409–434.

[64] G. SZEGÖ, *Orthogonal Polynomials*, vol. 23, American Mathematical Society Colloquium Publications, New York, revised ed., 1959.

[65] G. M. TREECE, R. W. PRAGER, AND A. H. GEE, *Regularised marching tetrahedra: Improved iso-surface extraction*, Computers and Graphics, 23 (1999), pp. 583–598.

[66] G. TURK AND J. F. O'BRIEN, *Shape transformation using variational implicit surfaces*, in SIGGRAPH'99, Aug 1999, pp. 335–342.

[67] ——, *Variational implicit surfaces*, Tech. Rep. GIT-GVU-99-15, Georgia Institute of Technology, May 1999.

[68] G. WAHBA, *Spline interpolation and smoothing on the sphere*, SIAM J. Sci. Stat. Comp, 2 (1981), pp. 5–16.

[69] H. WENDLAND, *Piecewise polynomial, positive definite and compactly supported radial basis functions of minimal degree*, Advances in Computational Mathematics, 4 (1995), pp. 389–396.

[70] ——, *Error estimates for interpolation by compactly supported radial basis functions of minimal degree*, J. Approx. Theory, 93 (1998), pp. 258–272.

[71] D. V. WIDDER, *The Laplace Transform*, Princeton University Press, Princeton, NJ, 1946.

[72] Z. WU, *Compactly supported positive definite radial functions*, Advances in Computational Mathematics, 4 (1995), pp. 283–292.

Appendix A

A table of inner and outer functions

$$T_0 = \begin{bmatrix} 1 \end{bmatrix}$$

$$T_1 = \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}$$

$$T_2 = \begin{bmatrix} z^2 & zw & w^2 \\ -2\,z\overline{w} & z\overline{z} - \overline{w}w & 2\,w\overline{z} \\ \overline{w}^2 & -\overline{w}\overline{z} & \overline{z}^2 \end{bmatrix}$$

$$T_3 = \begin{bmatrix} z^3 & z^2 w & z w^2 & w^3 \\ -3\,z^2\overline{w} & z^2\overline{z} - 2\,\overline{w}zw & 2\,zw\overline{z} - \overline{w}w^2 & 3\,w^2\overline{z} \\ 3\,z\overline{w}^2 & -2\,z\overline{w}\overline{z} + \overline{w}^2 w & z\overline{z}^2 - 2\,\overline{w}w\overline{z} & 3\,w\overline{z}^2 \\ -\overline{w}^3 & \overline{w}^2\overline{z} & -\overline{w}\overline{z}^2 & \overline{z}^3 \end{bmatrix}$$

$$T_4 = \begin{bmatrix} z^4 & z^3 w & z^2 w^2 & z w^3 & w^4 \\ -4\,z^3\overline{w} & z^3\overline{z} - 3\,\overline{w}z^2 w & 2\,z^2 w\overline{z} - 2\,\overline{w}zw^2 & 3\,zw^2\overline{z} - \overline{w}w^3 & 4\,w^3\overline{z} \\ 6\,z^2\overline{w}^2 & -3\,z^2\overline{w}\overline{z} + 3\,z\overline{w}^2 w & z^2\overline{z}^2 - 4\,z\overline{w}w\overline{z} + \overline{w}^2 w^2 & 3\,zw\overline{z}^2 - 3\,\overline{w}w^2\overline{z} & 6\,w^2\overline{z}^2 \\ -4\,z\overline{w}^3 & 3\,z\overline{w}^2\overline{z} - \overline{w}^3 w & -2\,z\overline{w}\overline{z}^2 + 2\,\overline{w}^2 w\overline{z} & z\overline{z}^3 - 3\,\overline{w}w\overline{z}^2 & 4\,w\overline{z}^3 \\ \overline{w}^4 & -\overline{w}^3\overline{z} & \overline{w}^2\overline{z}^2 & -\overline{w}\overline{z}^3 & \overline{z}^4 \end{bmatrix}$$

$$O_0 = \begin{bmatrix} \dfrac{1}{z\overline{z} + w\overline{w}} \end{bmatrix}$$

$$O_1 = \frac{1}{(z\overline{z} + w\overline{w})^2} \begin{bmatrix} \overline{z} & -w \\ \overline{w} & z \end{bmatrix}$$

$$O_2 = \frac{1}{(z\overline{z} + w\overline{w})^3} \begin{bmatrix} \overline{z}^2 & -\overline{z}w & w^2 \\ 2\,\overline{zw} & \overline{z}z - \overline{w}w & -2\,wz \\ \overline{w}^2 & \overline{w}z & z^2 \end{bmatrix}$$

$$O_3 = \frac{1}{(z\overline{z} + w\overline{w})^4} \begin{bmatrix} \overline{z}^3 & -\overline{z}^2w & \overline{z}w^2 & -w^3 \\ 3\,\overline{z}^2\overline{w} & z\overline{z}^2 - 2\,\overline{wz}w & -2\,\overline{z}wz + \overline{w}w^2 & 3\,w^2z \\ 3\,\overline{zw}^2 & 2\,\overline{zw}z - \overline{w}^2w & \overline{z}z^2 - 2\,\overline{w}wz & -3\,wz^2 \\ \overline{w}^3 & \overline{w}^2z & \overline{w}z^2 & z^3 \end{bmatrix}$$

$$O_4 = \frac{1}{(z\overline{z} + w\overline{w})^5} \times$$

$$\begin{bmatrix} \overline{z}^4 & -\overline{z}^3w & \overline{z}^2w^2 & -\overline{z}w^3 & w^4 \\ 4\,\overline{z}^3\overline{w} & z\overline{z}^3 - 3\,\overline{wz}^2w & -2\,z\overline{z}^2w + 2\,\overline{wz}w^2 & 3\,\overline{z}w^2z - \overline{w}w^3 & -4\,w^3z \\ 6\,\overline{z}^2\overline{w}^2 & 3\,z\overline{z}^2\overline{w} - 3\,\overline{zw}^2w & \overline{z}^2z^2 - 4\,\overline{zw}wz + \overline{w}^2w^2 & -3\,\overline{z}wz^2 + 3\,\overline{w}w^2z & 6\,w^2z^2 \\ 4\,\overline{zw}^3 & 3\,\overline{zw}^2z - \overline{w}^3w & 2\,\overline{zw}z^2 - 2\,\overline{w}^2wz & \overline{z}z^3 - 3\,\overline{w}wz^2 & -4\,wz^3 \\ \overline{w}^4 & \overline{w}^3z & \overline{w}^2z^2 & \overline{w}z^3 & z^4 \end{bmatrix}$$